RECHERCHE EN BREF

MISE AU POINT D'UN NOUVEAU MODÈLE D'EFFICACITÉ POUR LES ENSEIGNANTS

Leçons tirées de l'étude de l'AIERS sur le contexte de la salle de classe

DORIS W. RYAN

Cette étude a été effectuée dans le cadre d'un contrat conclu avec le ministère de l'Éducation de l'Ontario.
Les opinions qu'y expose l'auteure sont les siennes et ne coïncident pas forcément avec celles du ministère.

SeanConway, ministre
Bernard J. Shapiro, sous ministre

Ontario
Ministère de l'Éducation

87-04910/

MISE AU POINT D'UN NOUVEAU MODÈLE D'EFFICACITÉ POUR LES ENSEIGNANTS

Leçons tirées de l'étude de l'AIERS sur le contexte de la salle de classe

---

DORIS W. RYAN

Cette étude a été effectuée dans le cadre d'un contrat conclu avec le ministère de l'Éducation de l'Ontario.

Les opinions qu'y expose l'auteure sont les siennes et ne concordent pas forcément avec celles du ministère.

SeanConway, ministre
Bernard J. Shapiro, sous ministre

© IMPRIMEUR DE LA REINE POUR L'ONTARIO, 1987

On peut commander d'autres exemplaires de ce document à l'une des adresses suivantes :

Services des publications
L'Institut d'études pédagogiques
de l'Ontario
252, rue Bloor ouest
Toronto (Ontario)
M5S 1V6
(416) 926-4707

Les commandes de moins de 30,00 $
doivent être accompagnées d'un
chèque ou mandat de poste fait à
l'ordre de l'Institut d'études
pédagogiques de l'Ontario.

Services des publications du MSG
880, rue Bay, 5e étage
Toronto (Ontario)
M7A 1N8
(416) 965-6015
(Sans frais) 1-800-268-7540
(Sans frais si votre indicatif
régional est le 807) Demandez
le Zénith 67200 au ou à la
téléphoniste.

Les commandes doivent être
accompagnées d'un chèque ou
mandat de poste rédigé à
l'ordre du Trésorier de l'Ontario.

Contrat 1051
ONO 4497

Données de catalogage avant publication (Canada)

Ryan, Doris W., 1934-
Mise au point d'un nouveau modèle d'efficacité
pour les enseignants

(Recherche en bref)
Publié aussi en anglais sous le titre : Developing a
new model of teacher effectiveness.
Bibliogr.: p.
ISBN 0-7729-2671-9

1. Enseignement. I. Ontario. Ministère de
l'éducation. II. Titre. III. Collection.

LB1027.R9214 1987 371.1'02 C87-099633-9

TABLE DES MATIÈRES

Avant-propos v

I. Aperçu général de la recherche sur l'efficacité des enseignants 1

- La conception que l'efficacité des enseignants est liée à leurs traits de personnalité : études sur les traits des enseignants et études corrélant ces traits avec les résultats obtenus chez les élèves 5
- Études sur les ressources et les résultats 9
- L'efficacité des enseignants vue comme un ensemble de comportements : études sur les processus et résultats 11
- Conclusions 16

II. Aperçu général et résumé de l'étude de l'AIERS sur le contexte de la salle de classe 18

- Plan de l'étude 20
- Pays participants 22
- Variables et catégories 24
- Instruments 25
- Préparation et analyse des données 28
- Principales conclusions et généralisations tirées de l'étude sur le contexte de la salle de classe (ÉCSC) 33
- Deux analyses secondaires des données de l'ÉCSC 43
- Conclusions 48

III. Un nouveau modèle d'efficacité des enseignants fondé sur la conception que l'enseignement est un ensemble de tâches 50

- Remise en question de la validité des suppositions processus-résultats 51

- L'enseignement : un ensemble de tâches 55

- Cadre des principales tâches des enseignants 56

- Implications de la conception que l'enseignement est un ensemble de tâches 64

- L'utilisation des tâches des enseignants pour faire le pont entre la recherche et la pratique 68

- Résumé et conclusions 76

Bibliographie 80

Annexe A 87

AVANT-PROPOS

La présente monographie s'adresse principalement aux enseignants, aux administrateurs scolaires, aux responsables de l'élaboration de politiques d'éducation et aux agents de formation des enseignants. Elle a pour but de stimuler des discussions réalistes, à la fois sur les idées que nous nous faisons de ce qui rend certains enseignants plus efficaces que d'autres et sur les actes que nous posons et les décisions que nous prenons sous l'influence de ces idées. Voici comment s'articule cette monographie : (1) aperçu général de la recherche sur l'efficacité des enseignants, (2) aperçu général et résumé d'une étude internationale sur les relations entre les pratiques d'enseignement, les activités en classe et l'apprentissage cognitif et affectif des élèves, et (3) discussion sur la mise au point d'un nouveau modèle d'efficacité pour les enseignants, provenant des leçons que nous avons tirées de l'étude internationale. La première partie donne le contexte permettant de comprendre les idées qui ont sous-tendu l'étude de l'AIERS et la deuxième partie explique les raisons pour lesquelles mes collègues et moi en sommes venus à remettre ces idées en question.

La présente monographie n'est pas un rapport de recherche technique sur l'étude de l'AIERS sur le contexte de la salle de classe et de ce fait, elle ne reprend pas les analyses statistiques que l'on pourra se procurer ailleurs. L'équipe Ryan, Anderson et Shapiro rédige en ce moment un rapport complet sur l'étude internationale et ce rapport sera publié par Pergamon Press, au cours de 1987, espère-t-on. Les lecteurs intéressés souhaiteront peut-être consulter un numéro spécial du journal Evaluation in Education (Ryan and Anderson, 1984) ou obtenir le microfilm de la réponse de l'Ontario à l'étude de l'AIERS (Ryan et associés, 1984), auprès du ministère de l'Éducation de l'Ontario.

Le sommaire de l'étude de l'AIERS que nous présentons ici, de même que les généralisations qui en découlent, sont issus du travail acharné de plusieurs personnes, en particulier Lorin Anderson et Garret Mandeville, de l'Université de la Caroline du Sud et Norbert Sellin, de l'Université de Hambourg. Le modèle d'efficacité des enseignants dont traite la troisième partie a été mis au point grâce à la collaboration de Lorin Anderson et de Robert Burns (du Far West Laboratory for Educational Research and Development de San Francisco). À tous et à chacun de ces collègues, je désire exprimer ma profonde gratitude.

D. W. Ryan

I. APERÇU GÉNÉRAL DE LA RECHERCHE SUR L'EFFICACITÉ DES ENSEIGNANTS

Les débats de la recherche sur l'efficacité des enseignants remontent à la fin du XIX^{e} siècle, au moment de la publication de la première étude connue sur le sujet (Kratz, 1896; cité par Medley, 1982). On s'intéressa alors de plus en plus à la recherche dans ce domaine et cet intérêt continue aujourd'hui à être très marqué. Il y a plus de 20 ans, par exemple, Gage (1960) a identifié quelque 10 000 études consacrées à l'efficacité des enseignants. En examinant, même rapidement, la documentation la plus récente sur l'éducation, on se rend compte que la recherche sur l'efficacité des enseignants continue à attirer les ressources des chercheurs, des agences de financement et des organisations professionnelles (Doyle, 1978).

Pour quelles raisons ce type de recherche suscite-t-il tant d'intérêt, depuis aussi longtemps? Tout simplement parce que les enjeux sont très élevés. L'éducation est l'une des entreprises sociales les plus importantes, sinon la plus importante, compte tenu du nombre de personnes touchées, du temps consacré au processus éducatif et de la dépense de fonds publics qu'elle signifie (Walberg, 1986). L'affirmation d'Adam Smith, il y a deux siècles, demeure pertinente : l'éducation constitue un investissement social utile car la richesse des nations dépend des capacités de leurs habitants.

Il est peu probable que les parents et le public en général relient leurs préoccupations sur l'école et l'éducation à d'abstraites théories sur le capital humain. De plus, il se peut bien qu'ils ne soient pas partisans ou adeptes de la recherche sur l'efficacité des enseignants ou, du reste, de recherche dans tout autre domaine. Pourtant, ils <u>tiennent</u> à ce que les enfants reçoivent le meilleur enseignement possible dans les écoles publiques. Par les articles publiés dans la presse populaire, il est évident que bien des parents et des membres du public continuent à considérer d'un oeil critique l'école en général et les enseignants en particulier, et ce en dépit de tous les aspects qui, selon les éducateurs professionnels, ont bénéficié d'améliorations : exigences pour l'obtention d'un diplôme d'enseignement, formation des enseignants, programmes d'études et autres. Par exemple, on a examiné les sondages d'opinion annuels des Américains sur leurs écoles publiques et constaté que pendant treize années consécutives, les parents ont exprimé des réserves sur la qualité de l'enseignement (Elam, 1978; cité par Bridges, 1986). En Ontario, les sondages menés par l'IÉPO ont également révélé les préoccupations du public en ce qui a trait à la qualité de l'enseignement dans nos écoles (Livingstone et al., 1985).

Les professionnels de l'enseignement n'acceptent peut-être pas toutes les critiques émises par les parents et le public, mais partagent réellement le désir d'un enseignement de qualité dans nos écoles. Pour eux, la maximalisation de l'efficacité des enseignants est un but principal (Medley, 1982). Bien que les incompétents ne représentent vraisemblablement qu'une faible proportion du corps enseignant, ils nuisent à la réputation de la majorité formée d'enseignants compétents «surmenés, sous-rémunérés et sous-appréciés pour leurs efforts et leurs accomplissements» (Bridges, 1986, p. 3). Ni les administrateurs scolaires, ni les enseignants ne traitent à la légère les exigences des critiques de débarrasser les écoles des enseignants incompétents. Mais quelle est la nature de l'incompétence chez un enseignant ou une enseignante? Comment les administrateurs scolaires peuvent-ils distinguer les enseignants incompétents devant être remerciés de ceux qui ont besoin d'aide pour augmenter ou améliorer leur champ de compétences? L'ouvrage récent de Bridges (1986), The Incompetent Teacher, montre que ces questions restent problématiques. Les professeurs des facultés d'éducation s'intéressent également à l'amélioration du rendement des enseignants suffisamment compétents. Par conséquent, ils cherchent des réponses à deux questions supplémentaires : Quelles sont les caractéristiques d'un enseignement efficace? Comment les administrateurs scolaires, les responsables de la formation des enseignants et les organisations professionnelles soutiennent-ils les enseignants en général pour qu'ils améliorent leur efficacité à aider les élèves à apprendre? Avec raison, les praticiens de l'enseignement espèrent que les chercheurs les aideront à répondre à toutes ces questions.

Que se passe-t-il du côté des chercheurs eux-mêmes? Les professeurs d'université ont prouvé leur intérêt en concevant des projets de recherche et en cherchant des sources de financement pour leurs recherches et il ne faudrait pas croire qu'ils sont motivés uniquement (ou peut-être principalement) par le désir de satisfaire aux exigences du poste qu'ils occupent. Il est certain que depuis 1970, un nombre grandissant de chercheurs se perçoivent davantage comme des «agents d'amélioration» et non pas comme de simples «observateurs» du processus éducatif (voir à ce propos la terminologie de Gage, 1966). Doyle (1978) cite plusieurs études conçues spécialement dans le but de fournir des sources de contenu pour la formation des enseignants ainsi que des «outils» dont les enseignants pourront se servir pour améliorer leurs pratiques d'enseignement. Nous discuterons brièvement de ces études. Qu'il suffise de mentionner que, dans le monde de la recherche en éducation, un grand nombre de recherches sur l'efficacité des enseignants peuvent être placées sous la bannière de la «recherche appliquée» ou «axée sur l'amélioration» et qu'elles sont destinées à procurer aux praticiens des connaissances qui leur seront utiles.

Quelles connaissances avons-nous de ce qui rend certains enseignants plus efficaces que d'autres? Comment ces connaissances sont-elles utilisées en pratique? Avant de tenter de répondre à ces questions, il semble utile de définir certains termes utilisés au cours de la discussion. Nous avons emprunté ces définitions à Medley et nous allons le citer directement :

> Le terme «efficacité de l'enseignant ou enseignante» désigne les résultats obtenus par un enseignant ou une enseignante ou les progrès accomplis par ses élèves pour atteindre un but éducatif précis. L'une des implications de cette définition suppose que l'efficacité d'un enseignant ou d'une enseignante doit être définie et uniquement évaluée en termes de comportements des élèves et non en termes de comportements des enseignants. Pour cette raison, et parce que la quantité d'apprentissage des élèves est fortement influencée par des facteurs placés hors du contrôle de l'enseignant ou enseignante, l'efficacité des enseignants ne sera pas considérée comme une caractéristique stable de l'enseignant ou enseignante en tant qu'individu, mais plutôt comme le produit de l'interaction entre certains traits des enseignants et d'autres facteurs qui varient selon le contexte dans lequel l'enseignant ou l'enseignante travaille.
>
> On trouve la preuve de l'importance de ces facteurs contextuels dans les résultats des recherches comparant les progrès de deux séries de classes dirigées par le même enseignant ou la même enseignante (habituellement au cours d'années successives). On a établi que la corrélation entre les mesures de l'efficacité d'un même enseignant ou d'une même enseignante prises dans des classes différentes n'est que de 0,30 environ (Rosenshine, 1970; Veldman et Brophy, 1974). Ceci indique que la fiabilité d'une pareille mesure n'est que de 0,30 environ, ce qui signifie qu'en s'appuyant uniquement sur les progrès des élèves, on peut très bien évaluer qu'un enseignant ou une enseignante est très efficace une année donnée, puis constater qu'il ou elle l'est très peu l'année suivante.
>
> Les «expériences d'apprentissage des élèves» désignent les divers comportements des élèves pendant que l'apprentissage s'effectue. Elles ne sont pas une caractéristique de l'enseignant ou enseignante, mais ont beaucoup à voir avec le degré d'efficacité d'un enseignant ou d'une enseignante, étant donné que la quantité d'apprentissage d'un élève dépend de ce qu'il fait (les expériences qu'il a). Tout effet d'un enseignant ou d'une enseignante sur l'apprentissage d'un élève doit résulter d'un effet particulier que l'enseignant ou l'enseignante a sur les expériences d'apprentissage de l'élève.
>
> Le «rendement de l'enseignant ou enseignante» désigne le comportement d'un enseignant ou d'une enseignante chargé d'une classe (à la fois en classe et en dehors de la classe). Le rendement de l'enseignant ou enseignante diffère de l'efficacité d'un enseignant ou d'une enseignante. En effet, le rendement d'un enseignant ou d'une enseignante se définit en termes de comportements et d'actions tandis que son efficacité se définit par rapport à ce que les élèves font. Le rendement et l'efficacité se ressemblent car tous deux sont le produit de l'interaction entre certains traits de l'enseignant ou enseignante et le contexte d'enseignement. Le rendement d'un enseignant ou d'une enseignante s'emploie souvent comme base d'inférence de son efficacité.
>
> La «compétence de l'enseignant ou enseignante» désigne l'ensemble des connaissances, habiletés et convictions que possède l'enseignant

ou l'enseignante et qu'il apporte à son enseignement. La compétence d'un enseignant ou d'une enseignante diffère de son rendement et de son efficacité en ceci qu'il s'agit d'un trait stable de l'enseignant ou enseignante qui ne change pas de façon appréciable lorsque l'enseignant ou l'enseignante passe d'un contexte à un autre. Elle ressemble au rendement de l'enseignant ou enseignante en ceci qu'elle aussi s'emploie comme base d'inférence de son efficacité.

On conçoit la compétence de l'enseignant ou enseignante comme une forme de répertoire. On appellera «compétences» de l'enseignant ou enseignante les connaissances, habiletés et convictions faisant partie de son répertoire. Selon le but poursuivi, on définira les compétences en termes d'éléments très petits, telle «l'habileté d'emmobiner un projecteur de 16 mm» ou d'éléments très vastes, comme «l'habileté de mettre en pratique le modèle d'organisation préalable». On définit une compétence en fonction de ce qu'un enseignant ou une enseignante peut faire, sait, ou croit. «L'habileté de diriger une classe» peut être une compétence, «la capacité de diriger une classe» n'en sera pas une.

Notons que «compétence» est un nom collectif et qu'il n'est pas au pluriel, tandis que «compétences» est au pluriel. D'habitude, on considère qu'une compétence est une qualité présente ou absente et que la compétence est une question de degrés : plus un enseignant ou une enseignante a de compétences, plus il (elle) est compétent(e).

Les «traits préexistants de l'enseignant ou enseignante» désignent les connaissances, habiletés et convictions que l'on s'attend à retrouver chez un enseignant ou une enseignante débutant sa formation professionnelle. Il s'agit principalement de traits de personnalité stables comme l'intelligence générale ou un intérêt pour les enfants), traits que l'on estime utiles au bon rendement d'un enseignant ou d'une enseignante, mais qu'un programme de formation des enseignants ne peut pas et ne devrait pas essayer de créer chez les étudiants qui ne les possèdent pas déjà.

- Medley (1982, pp. 1894-95) (traduction libre)

Comme nous le verrons, les recherches entreprises à différentes périodes ont porté sur les différentes conceptions de l'efficacité des enseignants. De plus, les études ont eu différents impacts sur la politique et la pratique éducatives. Étant donné que la pratique tend à accuser du retard par rapport à la recherche, les plus récentes conceptions, hypothèses et conclusions issues de la recherche en éducation n'ont eu, jusqu'à présent, qu'une influence limitée. Par contre, les conceptions, hypothèses et conclusions des premières recherches sont toujours sous-jacentes à la pratique éducative et servent de base à la plupart des décisions prises sur l'efficacité des enseignants. Au dire de Medley (1982, p. 1894) : «Des conclusions de recherches nullement valables se sont immiscées dans les convictions des responsables des décisions prises en éducation et ont laissé leurs traces». Nous allons donc examiner ces premières recherches.

La conception que l'efficacité des enseignants est liée à leurs traits de personnalité : études sur les traits des enseignants et études corrélant ces traits avec les résultats obtenus chez les élèves

Pendant le premier tiers du siècle, un seul paradigme de recherche a dominé les recherches sur l'efficacité des enseignants. Son approche était fort simple : on demandait à des élèves des classes élémentaires ou du secondaire, ou à des éducateurs d'enseignants ou à des éducateurs, ou à quelque amalgame de ces groupes d'essayer de se souvenir du meilleur enseignant ou de la meilleure enseignante qu'ils aient jamais rencontré(e) personnellement et de consigner par écrit ce en quoi cet enseignant ou cette enseignante se distinguait des autres. Le chercheur assemblait ensuite ces descriptions, les comparait et s'en servait pour établir une liste de caractéristiques qui étaient censées distinguer les enseignants efficaces de ceux qui ne l'étaient pas. Cette liste servait ensuite à évaluer l'efficacité des enseignants ou à déterminer si les candidats à la profession avaient l'étoffe de l'enseignant ou enseignante «efficace».

On retrouva ce genre d'approche dans une série d'études menées sur plus de 35 ans. Les études se distinguaient les unes des autres principalement par la variété des «juges» mis à contribution et par la façon de produire la ou les listes (Medley, 1982). L'exemple le plus complet et le plus détaillé de ce genre de recherches est le Commonwealth Teacher Training Study (Charters et Waples, 1929), une étude à grande échelle qui a produit une quantité de listes de diverses longueurs. Voici quelques traits typiques figurant sur ces listes (les six premiers sur une liste de 25) : adaptabilité, prévenance, enthousiasme, bon jugement, honnêteté et magnétisme. Nulle part dans cette série d'études y a-t-il eu quelque tentative que ce soit de vérifier si les élèves des enseignants possédant ces traits apprenaient vraiment davantage que les élèves confiés à des enseignants qui, jugeait-on, en étaient démunis (Medley, 1982).

Deux suppositions implicites dans ces études initiales font, dans une certaine mesure, toujours sentir leurs effets encore aujourd'hui. Selon la première, toute personne ayant déjà fréquenté une école est capable de bien juger de l'efficacité des enseignants. Joyce et Weil (1972, p. 3) font remarquer que cette supposition provient du fait «qu'un bon enseignement est comme une belle oeuvre d'art : beaucoup de gens sont persuadés qu'ils savent le reconnaître d'un seul coup d'oeil, même s'ils ont de la difficulté à expliquer le fondement rationnel de leur jugement.» Medley (1982) note que presque tout observateur du travail d'un enseignant ou d'une enseignante se fait une forte idée de l'efficacité de cet enseignant ou cette enseignante, même si son observation n'a duré que quelques minutes; il s'agit d'un phénomène universel. Comme on l'a noté ci-dessus,

la valeur de ces idées en termes d'apprentissage par les élèves n'a pas été établie. Néanmoins, il faut accepter l'hypothèse que l'on peut juger un bon enseignement de façon subjective si l'on veut donner de la crédibilité à l'approche utilisée dans les premières études sur l'efficacité des enseignants.

La deuxième supposition veut que le fait d'être un bon enseignant ou une bonne enseignante soit inné et non acquis. Cela se sous-entend dans la nature des qualités dégagées de l'analyse des impressions des «juges» dans ces études. On pourrait ranger la plupart de ces traits avec ce que Medley appelle les «traits préexistants chez l'enseignant ou l'enseignante». En tant que tels, il ne s'agit pas de traits acquis par une formation pédagogique professionnelle. Il ne semble pas que les juges (élèves ou éducateurs professionnels) qui ont contribué à l'établissement de ces listes ou que les chercheurs qui les ont dressées, fassent une distinction entre ces traits de la personnalité des enseignants et les autres types de variables reliées à l'efficacité des enseignants. En conséquence, Medley conclut :

> Cette confusion parmi ces types de variables, en réalité très différents, qui présentent divers rapports avec l'efficacité que possède ou acquiert un enseignant ou une enseignante, non seulement transpire dans les premières recherches, mais continue d'influencer et de brouiller les décisions administratives prises encore aujourd'hui sur la foi de ces listes. Nombre d'agents de formation des enseignants et d'administrateurs scolaires continuent de fonder leurs décisions sur ces mêmes impressions subjectives (Medley, 1982, pp. 1895-96).
> (traduction libre)

Parmi les traits que Medley classerait dans les variables des «compétences» et du «rendement de l'enseignant ou enseignante» (ces deux types de variables pourraient, du moins en théorie, être développés par les programmes de formation des enseignants), quelques-uns commencèrent à apparaître sur les listes de qualités des enseignants issues des recherches menées au milieu des années mil neuf cent trente, au moment où l'intérêt pour la recherche sur les «traits» commençait à s'estomper. Cependant que la première supposition décrite ci-dessus avait toujours cours, la seconde se modifia jusqu'à un certain point, bien que les traits préexistants chez l'enseignant ou l'enseignante continuèrent à dominer les listes générées par les études.

À la fin de cette période, Hart mena une étude instructive (1936). Des élèves d'une école secondaire devaient d'abord faire une liste des traits de l'enseignant ou enseignante qu'ils avaient le mieux aimé. Les traits les plus fréquemment mentionnés dans la liste qui en découla comprenaient l'aptitude à l'enseignement (une des compétences de l'enseignant ou enseignante) et cinq autres traits (de personnalité) : dynamisme; cordialité; intérêt pour les élèves; impartialité; équité dans les corrections et l'attribution des notes. Ensuite,

on demanda aux élèves de déterminer si l'enseignant ou l'enseignante avec lequel ou laquelle ils avaient le plus appris était le même ou la même que celui ou celle qu'ils avaient dit avoir préféré(e). Ceux qui répondaient non étaient invités à décrire en quoi l'enseignant ou l'enseignante le ou la plus efficace différait de l'enseignant ou enseignante qu'ils avaient le mieux aimé(e). La liste qui en résulta présentait des différences remarquables par rapport à la première. On n'y retrouvait aucun trait préexistant chez l'enseignant ou l'enseignante, mais une variable du rendement d'un enseignant ou d'une enseignante (est exigeant(e) avec les élèves), deux des compétences d'un enseignant ou d'une enseignante (possède davantage d'aptitudes à l'enseignement et connaît mieux la matière enseignée), ainsi qu'un aspect de l'efficacité d'un enseignant ou d'une enseignante (réussit mieux à faire respecter la discipline), tel que nous avons défini ce concept précédemment. Bien sûr, la liste des traits de l'efficacité était encore basée sur des perceptions subjectives du rendement d'un enseignant ou d'une enseignante.

Le recours à des impressions subjectives pour prendre des décisions concernant l'efficacité des enseignants n'est pas le seul produit des premières études. L'autre produit, c'est la grille d'évaluation des enseignants qui apparut dans l'un des premiers rapports annuels de la National Society for the Study of Education (Boyce, 1915). Cette grille gagna rapidement en popularité. En 1930, un chercheur, Barr, établit une liste de 209 différentes grilles d'évaluation pour les enseignants et les étudiants en pédagogie. Règle générale, ces listes donnaient un certain nombre de traits semblables à ceux énumérés dans les listes décrites ci-dessus, lesquelles reflètent des opinions particulières sur ce qui fait l'efficacité d'un enseignant ou d'une enseignante. Des observateurs des salles de classe (habituellement des superviseurs) se servent de ces grilles pour évaluer à quel degré l'enseignant ou l'enseignante semble posséder chacun des traits examinés. D'habitude, ces évaluations sont consignées selon une échelle numérique (de un à cinq ou de un à dix, par exemple). Souvent, on fait le total ou la moyenne des points d'évaluation afin d'en arriver à un résultat général de l'efficacité de l'enseignant ou enseignante. De cette façon-là, l'efficacité d'un enseignant ou d'une enseignante est évaluée à partir de données sur son rendement dans une salle de classe.

Les chercheurs furent les premiers à utiliser ce nouvel outil car il leur offrait une méthode pratique pour quantifier les perceptions des observateurs du rendement en classe (Medley, 1982). Certains chercheurs, c'est à leur honneur, essayèrent de vérifier la validité des grilles d'évaluation en calculant les corrélations entre les mesures moyennes du progrès des élèves dans un échantil-

lon de classes et les évaluations des superviseurs des enseignants chargés de ces élèves. La première de ces corrélations date de 1935 (Barr) et on en fit plusieurs autres avant, approximativement, 1959. Les résultats étaient toujours les mêmes : il n'existait pas de corrélation entre les évaluations des superviseurs sur l'efficacité des enseignants et les mesures de l'efficacité des enseignants (en termes de résultats obtenus par les élèves à des examens visant à mesurer leurs progrès réels). Bridges (1986, p. 8) rapporte les conclusions représentatives suivantes, tirées d'études semblables :

> [...] les surintendants, les superviseurs et les directeurs d'école avaient tendance à donner une mauvaise évaluation à de bons enseignants ou de bonnes enseignantes et une bonne évaluation à de mauvais enseignants ou de mauvaises enseignantes (le critère pour définir un bon ou un mauvais enseignant ou une bonne ou mauvaise enseignante étant ici les progrès réels des élèves) [...] Il faudrait se garder d'accepter les évaluations des surintendants, des superviseurs et des directeurs d'école comme étant les seuls critères d'évaluation valables jusqu'à ce que les personnes occupant ces postes aient reçu une formation leur permettant d'assumer cette responsabilité (McCall et Krause, 1959).
>
> [...] on n'a pas suffisamment corrélé les évaluations fondées sur les observations des superviseurs et celles fondées sur les résultats réels obtenus par les élèves (Anderson, 1954).
>
> Quels que soient les progrès accomplis par les élèves grâce à la compétence de l'enseignant ou enseignante, les évaluations des superviseurs ne reflètent pas assez justement cette réalité (Jones, 1946).
>
> (traduction libre)

Bien qu'on n'ait encore publié aucune recherche prouvant que les évaluations des superviseurs constituent des indicateurs valables sur l'efficacité des enseignants, on continue d'utiliser la grille d'évaluation sans avoir beaucoup modifié la forme qu'elle présentait au moment de son introduction (Medley, 1982). L'utilisation de grilles ou d'échelles d'évaluation semblables comporte certains aspects inquiétants. Si les évaluations des superviseurs ne sont pas valables, comme les recherches semblent clairement l'indiquer, cela veut dire que l'on continue de prendre des décisions mal avisées en ce qui concerne : (1) les étudiants aptes à recevoir un brevet d'enseignement; (2) les candidats qui seront reconnus compétents (une pratique que viennent d'adopter certains états américains); (3) les enseignants qui recevront un brevet permanent d'enseignement; et (4) les enseignants qui ont besoin de formation en cours d'emploi ou même de supervision et d'évaluation continues aux fins d'un congédiement possible (Medley, 1982).

Un autre aspect, probablement tout aussi troublant, est que l'instrument de mesure mis à la disposition des agents de formation des enseignants et des administrateurs scolaires représentait jusqu'à récemment tout ce qu'il y avait de

plus moderne dans le domaine, malgré le fait qu'il datait de 1915! Cela s'explique par le fait que les chercheurs n'avaient pas, jusqu'à récemment, réussi à mettre au point des méthodes plus fiables et mieux adaptées pour reconnaître un bon enseignement par l'observation. En effet, pendant longtemps, les chercheurs acceptèrent aussi les évaluations des observateurs comme étant celles de leurs mandataires. Puisque les observations de ces mandataires servaient à des fins de recherche, on offrait à ces observateurs une formation relative à la définition des critères et une formation pour mieux juger ces critères (les traits énumérés dans la grille d'évaluation de l'efficacité des enseignants). De plus, dans certains cas, l'évaluation de l'efficacité d'un enseignant ou d'une enseignante corrélait avec les progrès réels accomplis par les élèves, lorsque les observateurs avaient reçu une formation (Rosenshine, 1971).

Un paradigme de recherche qui finit par être connu sous l'appellation présages-résultats découla des premières recherches sur les «traits» et de la mise au point d'échelles d'évaluation (Dunkin et Biddle, 1974). Les variables «présages» incluent d'habitude les traits des enseignants et leurs expériences de formation (lesquelles, on le suppose, représentent leurs compétences); les variables «résultats» sont habituellement représentées par une évaluation des élèves sous la forme d'examens, quoique les premières études de ce genre considéraient les observateurs comme des intermédiaires pour les chercheurs sur l'efficacité des enseignants. J'ai mentionné précédemment des études qui tentaient d'établir des liens entre les variables «présages» et le rendement des élèves (résultats). Contentons-nous d'ajouter à ce propos que ces études ont permis de recueillir bien peu d'informations fiables sur les relations entre les traits des enseignants et les résultats des élèves (Dunkin et Biddle, 1974). Voici une conclusion représentative à laquelle est parvenu un comité sur l'efficacité des enseignants, de l'American Educational Research Association :

> Un simple fait demeure : après plus de quarante ans de diverses recherches sur l'efficacité des enseignants, nous ne disposons toujours pas de critères fiables pour aider les surintendants scolaires à engager des enseignants ou à leur offrir un poste permanent, pour aider les organismes à accorder leur permanence aux enseignants ou pour faciliter la tâche des facultés d'éducation qui désirent planifier ou améliorer leurs programmes de formation des enseignants (AERA, 1953, p. 657). (traduction libre)

Études sur les ressources et les résultats

Bien que les études sur les ressources et les résultats représentent une tradition de recherches sur les effets de l'école plus que sur l'efficacité des enseignants en tant que telle, nous les avons incluses ici pour deux raisons.

Premièrement, leurs conclusions ont parfois servi à donner du poids à l'affirmation que les enseignants font peu ou pas de différence dans l'apprentissage des élèves, une affirmation diamétralement opposée à l'hypothèse que «l'enseignant ou l'enseignante fait toute la différence», qui sous-tend la recherche sur l'efficacité des enseignants. Deuxièmement, l'étude sur le contexte de la salle de classe dont nous parlerons dans la deuxième section de cette monographie fut menée par l'Association internationale pour l'évaluation du rendement scolaire (AIERS) et fut la première étude de l'AIERS qui n'utilisait pas l'approche des ressources et résultats. Nous parlerons des sondages de l'AIERS sur les progrès dans la prochaine section.

L'une des études sur les ressources et résultats ayant reçu le plus de publicité est le sondage Equality of Educational Opportunity, de Coleman et al. (1966). Barr et Dreeben (1978) observèrent que ce sondage «a marqué le grand tournant de la recherche sur les effets de l'école», et a favorisé une expansion rapide de la recherche fondée sur le modèle de la production-fonction, emprunté au domaine de l'économie. Dans ce genre de recherches, on considère l'école comme une entreprise qui transforme des ressources (telles que des livres, des traits des enseignants, des activités, des installations physiques et de l'équipement) en résultats (tels que des connaissances accrues, des changements d'attitudes ou des progrès réels).

L'étude Coleman a été effectuée aux É.-U., dans 4 000 écoles publiques auprès de 645 000 élèves. Ses conclusions ont été interprétées (mal interprétées, selon Barr et Dreeben, 1978) comme prouvant que l'école n'avait que peu d'effets sur l'apprentissage des élèves. Les critiques de l'étude ont cependant fait remarquer que l'on n'avait examiné aucune variable sur le processus en salle de classe et que l'on n'avait pas non plus fait d'observations en classe ou à l'intérieur des écoles. Comme l'ont fait ressortir Averch et ses associés (1972, p. 4), «La recherche utilisant cette approche voit l'école comme une boîte noire qui contient des élèves... des ressources sont appliquées aux élèves dans la boîte et il en découle des résultats». Bref, le modèle de production-fonction ne tient pas compte de ce que «l'entreprise» fait pour transformer des matières premières en résultats (Barr et Dreeben, 1978). Une autre critique du même ordre fut que l'étude considérait le rendement des élèves en fonction des différences entre les écoles plutôt qu'en fonction des différences entre les enseignants ou les classes. Autrement dit, les traits de l'enseignant ou enseignante, de l'école et des élèves furent tous réunis au niveau de l'école. Cependant, Good, Biddle et Brophy (1975) ont souligné que «si les enseignants font une différence, et que si les élèves sont exposés à la fois à du bon enseigne-

ment et à du mauvais enseignement, il ne devrait pas être surprenant que plusieurs écoles aient des effets fort semblables sur leurs élèves».

Malgré les critiques à cette approche et le manque de recherches décrivant ce qui se passe en classe, les études des années soixante sur les ressources et les résultats servirent à appuyer l'idée que l'efficacité des enseignants n'est pas différentielle. Selon cette idée, les élèves apprennent en raison de variables n'ayant que peu ou rien à voir avec les enseignants ou le milieu de la salle de classe, c'est à dire : aptitudes des élèves, intérêt pour l'instruction, degré d'encouragement de la part des parents (voir Stephens, 1967). Good, Grouws et Ebmeier (1983, pp. 3-4) ont affirmé que :

> ...à la fin des années soixante et au début des années soixante-dix, c'était presque devenu un sport national que de critiquer les enseignants et l'école. Des livres populaires, comme Death at an Early Age (Kozol, 1967) ou Crisis in the Classroom (Silberman, 1970), parmi d'autres, étaient souvent donnés comme preuve que l'école ou les enseignants ne faisaient que peu de différence dans les progrès des élèves. (traduction libre)

Ces auteurs ont fait également valoir que l'un des arguments implicites dans toute cette documentation est qu'il ne semblait pas y avoir beaucoup de variations dans la routine des écoles et des classes; «quand on voyait une classe, on les voyait toutes.» Ils soutenaient que les défenseurs de la conclusion «aucune différence» devraient prouver que (1) tous les enseignants enseignent de la même façon ou (2) que les différences dans l'enseignement aboutissent toutes aux mêmes résultats chez les élèves.

C'est dans ce contexte que le paradigme de recherche «ressources et résultats» en vint à servir de «cadre organisationnel à la plupart des recherches contemporaines sur l'efficacité de l'enseignement» (Doyle, 1978, p. 165). L'étude de l'AIERS sur le contexte de la salle de classe s'est appuyée sur ce modèle de recherche que nous allons maintenant aborder.

L'efficacité des enseignants vue comme un ensemble de comportements : études sur les processus et résultats

Un article d'encyclopédie par Mitzel (1960) a marqué le début d'un changement important dans la recherche sur l'efficacité des enseignants. Mitzel a défini trois catégories de critères d'efficacité des enseignants : le critère des présages (traits et compétences de l'enseignant ou enseignante), le critère des processus (les variables dans le rendement de l'enseignant ou enseignante) et

le critère des résultats (les résultats obtenus par les enseignants ou les progrès accomplis par les élèves vers un but pédagogique particulier, c'est à dire la définition de l'efficacité des enseignants mentionnée précédemment dans la présente section). Il a également ajouté une quatrième catégorie de variables liées à l'efficacité des enseignants qu'il a appelées les variables «contextuelles». Celles-ci comprenaient toute une série de facteurs contextuels ayant une incidence sur l'apprentissage des élèves, mais sur lesquels l'enseignant ou l'enseignante n'a pas de contrôle. Medley (1982, p. 1896) a affirmé que l'article de Mitzel «semble avoir été le début d'un processus de différenciation de ces variables, qui permettra aux chercheurs de les étudier séparément».

Peu après, Gage (1963) a écrit un article sur un important développement : l'introduction de plans d'observation à «faible inférence» comme une autre solution aux échelles d'évaluation pour juger du rendement des enseignants. Ces plans sont dits à «faible inférence» car ils prévoient que l'observateur fera relativement peu d'inférences ou d'extrapolations sur ce qui est objectivement visible ou audible dans la salle de classe. L'article de Gage décrivait aussi une méthode d'étude de l'efficacité des enseignants qui en vint à être appelée recherche «processus-résultats». Voici comment cela se présentait : (1) on observait des échantillons du travail de chacun des enseignants d'un groupe donné et on notait la récurrence de tel ou tel comportement en fonction de catégories prédéfinies (à l'occasion, on se servait aussi d'une échelle d'évaluation pour mesurer certains aspects du travail des enseignants, qui étaient difficiles à définir en termes de catégories précises de comportements); (2) on obtenait une mesure de l'efficacité de chacun des enseignants en ajustant les résultats moyens de la classe à un post-test, choisi pour mesurer les connaissances acquises de façon à tenir compte des différences dans les résultats obtenus à un prétest (en utilisant l'analyse de covariance); et (3) on calculait les corrélations entre les mesures sur le rendement des enseignants (la récurrence moyenne des comportements désignés, à l'échelle de toutes les classes observées, et pour chacun des enseignants) et les mesures sur le rendement moyen des classes.

Au cours des années soixante-dix aux États-Unis, on vit un grand nombre d'études qui suivaient la voie du paradigme des processus-résultats. Crawford et Gage (1977) passèrent quatre de ces études en revue; chacune des quatre était une étude à échelle relativement grande de classes élémentaires (1ère, 2e et 3e années) et on y appliquait une variable «résultats» pour mesurer le rendement en lecture des élèves. Il s'agit des études de Brophy et Evertson (1974), McDonald et Elias (1976), Soar (1973), et Stallings et Kaskowitz (1974). Les inférences (les affirmations débutant par «l'enseignant ou l'enseignante

devrait...») qu'ils en tiraient étaient les suivantes :

> Les enseignants devraient adopter un système de règles qui permette aux élèves de satisfaire à leurs besoins personnels et méthodologiques sans avoir à obtenir l'approbation de l'enseignant ou enseignante.
>
> Les enseignants devraient se déplacer beaucoup dans la salle de classe en surveillant le travail individuel des élèves et en leur laissant savoir qu'ils sont bien conscients de leurs comportements, tout en répondant également à leurs besoins éducatifs.
>
> Quand les élèves font du travail individuel, les enseignants devraient s'assurer que les devoirs sont intéressants, utiles et en même temps assez faciles pour que chaque élève de 3e année puisse les faire au complet sans l'aide de l'enseignant ou enseignante.
>
> Les enseignants devraient limiter le plus possible les activités d'organisation des cours. Il suffit qu'ils indiquent les activités de la journée au tableau, afin que les élèves sachent à quoi s'en tenir, où se rendre, etc.
>
> En choisissant un élève pour répondre à une question, les enseignants devraient l'appeler par son nom avant de poser la question. Il s'agit d'une façon de s'assurer que tous les élèves ont des chances égales de répondre aux questions.
>
> Dans le cas des élèves qui sont moins portés vers le travail scolaire, les enseignants devraient toujours tenter d'obtenir quelque élément de réponse aux questions. Reformuler la question, donner des indices ou encore poser une autre question, voilà des techniques utiles pour aider un élève qui garde le silence, qui répond «je ne sais pas» ou qui donne une mauvaise réponse.
>
> Au cours de sessions d'apprentissage de la lecture en groupe, les enseignants devraient donner le plus souvent possible des encouragements brefs et prévoir des exercices au rythme rapide.
>
> - Gage (1978, p. 39) (traduction libre)

Il est intéressant de comparer ces inférences avec celles provenant d'un compte rendu de Medley (1977) sur quatorze études, à grande échelle pour la plupart, et subventionnées par le gouvernement fédéral, de classes élémentaires regroupant une majorité d'élèves issus d'une catégorie socio-économique faible. Les «résultats» étaient ceux de tests de mathématiques ou de lecture. Medley n'examina que les corrélations qui se retrouvaient dans deux études différentes ou plus. Ces corrélations «fiables» concernaient trois aspects du rendement des enseignants en salle de classe. Le premier de ces aspects était le climat d'apprentissage que l'enseignant ou l'enseignante établissait et maintenait. Les classes où les enseignants réussissaient le mieux étaient celles où l'ordre régnait, où l'enseignant ou l'enseignante était plus sévère, où il ou elle consacrait le

moins de temps à la «direction», et où les élèves recevaient le plus d'encouragements et le moins de remontrances. Le second aspect était l'utilisation que l'enseignant ou l'enseignante faisait du temps des élèves. Les élèves des enseignants les plus efficaces passaient davantage de temps à des activités scolaires, et plus de temps en grand groupe pris en charge par l'enseignant ou l'enseignante; ils faisaient moins de travail en groupes restreints ou de travail individuel en classe et les enseignants les supervisaient de façon très soutenue. Ces résultats ont mené à la définition d'une variable appelée «temps réservé à l'apprentissage» (Rosenshine et Berliner, 1978) ou «temps d'apprentissage scolaire» (Fisher et al., 1980), une variable en étroite relation avec l'apprentissage des élèves. Finalement, le troisième aspect était la stratégie que suivaient les enseignants au cours des discussions en classe. Les enseignants les plus efficaces posaient plus de questions et les élèves en posaient moins; les enseignants posaient moins de questions de «niveau cognitif élevé» et plus de questions de «niveau cognitif faible»; et l'enseignant ou l'enseignante était moins porté(e) à réagir aux réponses des élèves, à les développer ou à les commenter. Ainsi, la stratégie préférée des enseignants les plus efficaces ressemblait à des exercices et à des récitations bien plus qu'à ce qu'on associe normalement à des discussions.

Les conclusions d'un troisième compte rendu portant celui-là sur un nombre d'études encore plus grand ont été rapportées par Rosenshine (1976). Il organisa ses conclusions en une stratégie, ou modèle d'enseignement qu'il appela «l'enseignement direct». Ce concept désigne des classes orientées vers l'acquisition de la matière, dirigées par l'enseignant ou l'enseignante et faisant appel à des matières structurées et présentées dans un ordre déterminé. En utilisant les techniques d'enseignement direct, le but d'un enseignant ou d'une enseignante est d'agir sur les élèves par le biais d'une suite organisée de matières ou de tâches en s'assurant : que les élèves comprennent clairement les objectifs visés; que le temps alloué pour l'apprentissage soit suffisant et continu; que la matière soit entièrement couverte; que le rendement des élèves soit bien contrôlé; que les questions soient de niveau cognitif faible de manière à ce que les élèves répondent souvent correctement; et que les réactions de l'enseignant ou enseignante aux réponses et interventions des élèves soient immédiates et axées sur la matière (Rosenshine, 1979).

Ce qui est encore plus important, c'est que les trois comptes rendus principaux s'accordaient tous pour dire que la recherche sur les processus et les résultats des années soixante-dix avait donné naissance à des notions utiles et s'accordaient également de façon assez unanime sur la nature de ces nouvelles notions.

Bien que ces découvertes aient été limitées au niveau élémentaire et, très souvent, à des enfants d'une clientèle dite «défavorisée», elles entraient en conflit avec l'image de l'enseignant ou enseignante efficace que les programmes de formation des enseignants présentent généralement à leurs étudiants. Au dire de Medley (1982, p. 1898) : «Si on continue, comme par le passé, à conseiller aux étudiants au baccalauréat en sciences pédagogiques d'adopter une attitude tolérante, de faire travailler les élèves en groupes restreints autonomes et d'éviter l'enseignement en grand groupe, alors il est temps que les responsables de la formation des enseignants reconsidèrent d'un oeil très critique les recherches qui ont donné cours à ces idées». (traduction libre)

Toute convaincante qu'ait pu être l'uniformité des conclusions émises par ces études corrélationnelles, les relations de cause à effet devaient être vérifiées par des études expérimentales dans lesquelles les enseignants recevaient une formation visant à leur faire adopter les comportements attribués aux «bons» enseignants et dans lesquelles on mesurait les effets sur l'apprentissage des élèves. Déjà en 1979, au moment de la publication de l'étude sur le contexte de la salle de classe de l'AIERS, quatre études expérimentales du genre avaient été publiées. On les devait à Anderson, Evertson, et Brophy (1979), Crawford et al. (1978), Good et Grouws (1979), et Stallings, Needels et Stayrook (1979). En général, ces études confirmaient les résultats des études corrélationnelles dont dérivaient les «traitements» de la formation des enseignants. De façon plus importante, elles prouvaient que les progrès des élèves augmentent considérablement lorsque les enseignants reçoivent une formation visant à leur faire adopter ces nouvelles idées (Gage et Fordham, 1980).

Les études sur les processus et les résultats des années soixante-dix ont contribué de façon remarquable à augmenter les connaissances sur l'efficacité des enseignants. Elles ont fourni les «premières découvertes claires et applicables en quatre-vingts ans de recherche sur l'efficacité des enseignants» (Medley, 1982, p. 1898). En effet, Medley a conclu en 1979 que les conclusions et les implications de ces études sur les salles de classe où l'enseignement est axé sur les tâches, le temps d'apprentissage scolaire, et les processus d'enseignement direct devraient être obligatoires dans les programmes de formation des enseignants et pour les enseignants en fonction (Medley, 1979).

Néanmoins, presque avant la publication des premiers résultats de ces recherches, le paradigme des processus-résultats fut l'objet de critiques plutôt acerbes, qui prirent plusieurs formes (voir Gage, 1978, pour une revue et une réfutation). Plusieurs de ces critiques concernaient les méthodes d'observation

des comportements des enseignants et de traitement des données (e.g. réunir les données recueillies durant tous les cours observés). Nous discuterons de ces critiques dans la troisième partie de la présente monographie car les expériences que nous avons acquises grâce à l'étude de l'AIERS illustrent nombre d'arguments des critiques.

Cependant, il convient de noter ici que le recours à l'observation en classe et l'utilisation de listes de vérification de comportements par les administrateurs scolaires qui supervisaient et évaluaient les enseignants sont devenus assez répandus vers la fin des années soixante-dix. On peut attribuer cette tendance aux études sur les processus et les résultats et à leur utilisation de systèmes d'observation à niveau d'inférence faible. Ces outils de recherche ne conviennent pas à la supervision des enseignants, un point que nous avons tenté de faire valoir lorsque nous avons pris la parole à des ateliers sur le perfectionnement des enseignants (Ryan, 1981).

Conclusions

L'examen de la documentation présentée ci-dessus s'est limité aux deux principales «écoles de pensée» qui ont guidé la recherche sur l'efficacité des enseignants depuis le tournant du siècle jusqu'à la fin des années soixante-dix. Nous les avons appelées les «traits de l'enseignant ou enseignante» et les «comportements de l'enseignant ou enseignante». Les deux écoles ont eu de l'impact sur les pratiques éducatives, mais il semble que l'influence de la perception des «traits» ait été plus envahissante, malgré les doutes sérieux qu'a suscités sa validité en tant que méthode employée pour déterminer ce qui rend un enseignant ou une enseignante plus efficace qu'un autre ou qu'une autre dans sa tâche qui est de favoriser l'apprentissage des élèves. On a également émis des réserves sur la validité de la perception que l'enseignement est un ensemble de comportements, comme nous le verrons. Cependant, les chercheurs tenants de la théorie des processus et des résultats qui ont fondé leurs recherches sur la perception que l'enseignement est une affaire de comportements ont véritablement réussi à exprimer des implications générales (ou des applications de leurs découvertes) qui semblent en même temps valables pour prévoir les progrès des élèves et utiles pour améliorer la pratique de l'enseignement.

L'esprit d'optimisme qui a prévalu chez les chercheurs à la fin des années soixante-dix n'a pas beaucoup souffert des critiques de plus en plus nombreuses qui se sont abattues sur la perception de l'enseignement qui était implicite dans le modèle des processus et des résultats. C'est dans ce climat d'optimisme

que fut conçue et menée l'étude sur le contexte de la salle de classe de l'AIERS. Nous allons maintenant aborder une discussion de cette étude.

L'étude de l'AIERS a commencé juste au moment où l'on améliorait les prémisses des premières études fondées sur les processus et les résultats. De plus, de nouvelles écoles de pensée émergèrent et dictèrent quelles étaient les meilleures façons de mener la recherche en salle de classe. Les résultats de ces études «processus-résultats» améliorées et des études fondées sur différents points de départ n'étaient pas tous disponibles lors de la conception de l'étude de l'AIERS. Ils font effectivement partie des leçons que nous avons tirées en nous efforçant de comprendre tous les éléments que nous avons rassemblés.

II. APERÇU GÉNÉRAL ET RÉSUMÉ DE L'ÉTUDE DE L'AIERS SUR LE CONTEXTE DE LA SALLE DE CLASSE

L'Association internationale pour l'évaluation du rendement scolaire (AIERS), un consortium international des instituts de recherche, fut fondée au cours de la période des études basées sur le modèle des «ressources-résultats» à grande échelle des années soixante. Comme nous l'avons déjà vu, les conclusions de plusieurs de ces études incitaient à penser que le rendement des élèves est influencé principalement par des facteurs et des variables extérieurs à l'école et à la salle de classe. Par exemple, les conclusions d'une étude pilote menée en 1961-1962, juste avant la première étude comparative de l'AIERS sur le rendement en mathématiques, montra que le rôle des «dispositions innées» dans l'explication des variances notées dans la résolution d'une question de test était plus important que «l'insistance en classe» d'un enseignant ou d'une enseignante (Purves, sous presse). Le rapport Equality of Educational Opportunity (Coleman et al., 1966) faisait état de conclusions semblables, mais en ayant substitué les dispositions innées par les antécédents familiaux, et remplacé «l'insistance en classe» des enseignants par les installations physiques, les programmes formels et les traits des enseignants.

Les premières études de l'AIERS étaient des études comparatives du rendement dans diverses matières. En 1970, on commença à disposer des premières conclusions de ces études. Ces dernières donnaient à penser que seulement 40 % à 50 % des différences dans le rendement des élèves pouvaient s'expliquer par les variables «ressources» utilisées par les modèles internationaux (se référer à Thorndike, 1983; Comber et Keeves, 1973; Purves, 1973). À l'AIERS, on s'inquiéta, avant même la publication des conclusions, de deux faiblesses dans la conception des études : (1) ces études étaient transversales plutôt que longitudinales et (2) elles ne donnaient pas de mesures de l'enseignement réel. Il faut se souvenir que l'on n'en savait pas beaucoup à l'époque sur ce qui se passait dans les salles de classe car peu d'études comportaient des phases d'observation en salle de classe. Dans les études ressources-résultats, les écoles et leurs salles de classe étaient conçues comme des «boîtes noires». En conséquence, Anderson (sous presse) conclut : «Les tentatives de les caractériser a résulté en affirmations qui équivalaient à dire que les écoles et les classes étaient des "boîtes" et qu'elles étaient "noires"».

Plusieurs des chercheurs de l'AIERS partageaient l'opinion que si les études

longitudinales pouvaient être conçues de façon à mesurer le rendement à deux reprises ou plus au cours d'une période donnée, et que si les comportements des enseignants étaient observés au cours de la ou des périodes de temps comprises entre les mesures du rendement, on pourrait alors expliquer la variation supplémentaire dans le rendement des élèves. De plus, des modèles d'étude semblables pourraient permettre d'en savoir plus long sur les facteurs reliés à l'école qui contribuent aux différences dans l'apprentissage des élèves. Cependant, au sein de l'AIERS, des chercheurs de différents pays n'étaient pas d'accord sur la faisabilité ou l'utilité d'études aussi exigeantes. Ainsi, bien qu'on ait tenu à Stockholm en 1972 une assemblée de chercheurs intéressés et qu'on ait préparé la proposition d'une éventuelle étude internationale, le projet resta lettre morte pendant de nombreuses années. Au cours de ces années, comme nous l'avons vu, un nombre grandissant d'études sur l'enseignement en classe furent menées aux É.-U. et les conclusions compilées des études publiées à la fin des années soixante-dix suggéraient qu'il existait des «principes d'enseignement» (Anderson, Evertson et Brophy, 1979), des «pratiques d'enseignement» (Crawford et al., 1978) ainsi que des «profils de comportement» (Good et Grouws, 1979) qui étaient associés au rendement des élèves et qui, vraisemblablement, influençaient ce dernier. Donc, il semblait à la fois possible et souhaitable de tenter d'effectuer une étude semblable au niveau international.

En 1978, un petit groupe de chercheurs de l'AIERS se réunirent en Israël pour discuter des divers modèles de recherches ou d'approches pouvant être envisagés. L'ébauche du projet retenu à cette réunion (Levin et Lewy, 1978) suggérait une étude expérimentale. Ces chercheurs proposèrent qu'un groupe d'enseignants de chaque pays reçoive une formation sur les variables des processus en salle de classe qui semblaient, selon les études antérieures, entrer en ligne de compte dans l'apprentissage des élèves. Dans chaque pays, un groupe de contrôle ne recevrait pas cette formation spéciale. Ensuite, on comparerait les deux groupes d'enseignants et leurs classes, selon les processus observés en salle de classe et les résultats moyens du test de rendement de leurs élèves. Les variables «ressources» utilisées dans le programme de formation comprenaient les cinq groupes suivants : temps d'apprentissage, questions en classe, indices d'apprentissage et d'enseignement, réactions et correctifs, et climat d'apprentissage. Étant donné que ces groupes avaient été identifiés à partir d'analyses d'autres études, la proposition suggéra que l'on procède à l'expérience sans avoir recours à une étude corrélationnelle.

L'AIERS organisa par la suite deux réunions pour discuter du projet Levin-Lewy. La première de ces réunions eut lieu à Berlin en 1979 et la seconde à Hambourg

en 1980. La portée recherchée et la nature de l'étude internationale proposée changèrent considérablement au cours des discussions tenues au cours de ces deux réunions. On avait beaucoup de réticence à s'engager dans une étude expérimentale fondée sur les comportements des enseignants et les variables «processus» identifiées principalement par les études américaines. On fit remarquer que rien n'incitait à croire que les processus en salle de classe dans les différents pays concordaient (ou, du reste, s'opposaient). Il semblait y avoir trop d'inconnues pour justifier une étude expérimentale avant d'avoir tenté et réussi une première étude descriptive. C'est la raison pour laquelle le groupe convint de suivre un modèle «processus-résultats» en deux volets, tel que le modèle américain, à savoir une étude corrélationnelle «naturelle» suivie d'une étude expérimentale.

Bref, l'étude de l'AIERS sur le contexte de la salle de classe (ÉCSC) fut conçue pendant une période d'optimisme renouvelé en ce qui concernait l'impact potentiel de l'école et de la salle de classe sur le rendement des élèves. L'esprit de ces cheminements fut bien rendu par un participant australien de l'étude originale :

> L'identification des comportements des enseignants qui influençaient le plus le rendement des élèves et leurs attitudes a été au coeur des préoccupations des chercheurs en éducation au cours des dernières années. Le pessimisme qui avait d'abord entouré la plupart des efforts des chercheurs dans ce domaine a été remplacé, au début des années soixante-dix, par le sentiment qu'il existait dorénavant une somme substantielle de connaissances au sujet de l'influence des pratiques éducatives sur l'apprentissage des élèves (Fordham, 1983).
> (traduction libre)

Le professeur N. L. Gage de Stanford fut le premier président du conseil responsable de l'étude internationale et Tamar Levin, de l'Université de Tel Aviv, en devint le coordonnateur. Pour un certain nombre de raisons, ils démissionnèrent tous les deux de leur poste avant que ne commence la mise en application de l'étude. On choisit pour les remplacer Bernard Shapiro, comme président et Doris Ryan, comme coordonnatrice, tous deux de l'IÉPO.

Plan de l'étude

L'étude de l'AIERS sur le contexte de la salle de classe comportait deux buts principaux. Premièrement, elle devait identifier les ressemblances et les différences dans les pratiques éducatives et les conditions d'apprentissage en vigueur dans une variété de pays. Deuxièmement, les résultats de l'étude descriptive allaient servir à identifier les comportements des enseignants qui

s'assortissaient d'un accroissement du rendement chez les élèves. Dans la mesure où il était possible d'identifier ces comportements, on allait les réunir en pratiques d'enseignement plus générales qui pourraient être apprises par les enseignants dans le cadre d'une étude expérimentale, en vue d'améliorer et l'enseignement et l'apprentissage. En réalité, seule l'étude corrélationnelle fut menée et les discussions qui suivent se limitent à la phase un de l'étude descriptive.

Au départ, l'étude corrélationnelle devait être longitudinale. Chacun des centres nationaux participants devait choisir une ou plusieurs unités d'apprentissage nécessitant une période d'apprentissage déterminée. Afin de décider de la période de temps devant s'écouler entre l'administration d'un prétest au début d'une unité et l'administration d'un post-test à la fin d'une unité, on allait se fier au nombre et au type d'unités choisies par chaque centre national. Dans des écoles anglophones de l'Ontario, par exemple, le prétest eut lieu au début de l'année scolaire et le post-test, à la fin de l'année; de cette manière, on couvrait le programme de mathématiques de la 8e année au complet. Par contre, en République de Corée, on choisit pour l'étude une simple unité (une matière), les sciences en 5e année; ainsi, seulement deux mois séparèrent le prétest et le post-test. Dans l'intervalle entre le prétest et le post-test dans chaque pays, on devait observer de six à dix fois, à périodes fixes, chaque classe de l'échantillon. Ces observations avaient pour but de recueillir des données sur les comportements d'enseignement et de direction des enseignants et des données sur la participation des élèves aux activités éducatives. On devait recueillir, à différents moments, avant et après les observations en classe, des données sur les caractéristiques de la collectivité, de l'école, des classes, des parents, des enseignants et des élèves. Enfin, on devait procéder à des analyses visant à déterminer la nature et la portée des relations entre les pratiques éducatives observées et le rendement et les attitudes des élèves dans le contexte des caractéristiques de la collectivité, de l'école, des classes, des parents et des élèves.

Même si tous les pays participants ont suivi le concept de base décrit ci-dessus, ils étaient libres de faire un certain nombre de choix importants. D'abord, chaque centre national pouvait choisir la matière, l'âge, le degré scolaire des élèves visés par leur version de l'étude. En raison de ce choix, et contrairement à toutes les autres études de l'AIERS, chaque centre national pouvait également décider de concevoir ou de choisir le prétest et le post-test utilisés dans l'étude pour évaluer le rendement des élèves. Comme nous l'avons déjà noté, les pays pouvaient aussi déterminer la quantité de matière à couvrir entre

le prétest et le post-test. Cependant, malgré ces différences au niveau des pays, chacune des versions de l'étude de l'AIERS avait pour but de recueillir des données sur un ensemble déterminé de variables en utilisant au moins douze instruments mis au point au niveau international.

Pays participants

Neuf pays participèrent à l'étude et le Canada fournit trois versions. L'Ontario dirigea une version dans ses écoles anglophones et une autre dans ses écoles francophones. Le Québec participa avec une étude dans ses écoles francophones. Cependant, aux fins des analyses internationales, on considéra chacune de ces études comme provenant de «pays» différents. En plus du Canada, les autres pays participants comprenaient l'Australie, les Pays-Bas, la République fédérale d'Allemagne, la Hongrie, Israël, la République de Corée, le Nigéria et la Thaïlande. Ces pays, situés sur cinq continents, permirent à l'étude internationale de recueillir une base de données de diverses cultures.

Le tableau 1 présente des renseignements sur les études menées dans les pays participants à l'exception de l'Allemagne fédérale. Les études de tous les pays, à l'exception de celle de la RFA furent menées entre 1980 et 1983; l'étude allemande fut menée au cours de 1984, trop tard pour qu'on puisse l'inclure dans les analyses de données internationales.

Comme l'indique le tableau 1, deux pays choisirent les mathématiques en 5^{e} année, cinq pays optèrent pour les mathématiques en 8^{e} année. Les autres choisirent la physique en 5^{e}, les sciences générales en 5^{e} et l'histoire ancienne en 6^{e}. Les lignes directrices de l'étude recommandaient un minimum de 40 classes d'élèves par pays, mais tous les pays ne purent adhérer à cet échantillon cible. Selon les versions, le nombre des classes variait de 18 à 87; donc, le nombre d'élèves par pays fluctuait aussi considérablement, de 400 à plus de 2 500. La durée de l'étude (c'est à dire le temps écoulé entre le prétest et le post-test) s'étala entre deux mois et neuf mois. Enfin, la majorité des pays procédèrent à huit observations par salle de classe. Certaines salles de classe dans un des pays ne furent observées qu'une fois tandis que d'autres pays se livrèrent à dix observations par salle de classe.

TABLEAU 1

APERÇU GÉNÉRAL DE LA RÉALISATION DE L'ÉTUDE SUR LE CONTEXTE DE LA SALLE DE CLASSE

Pays	Année Scolaire	Matière	Nombre d'écoles	Nombre de classes	Nombre d'élèves	Durée de l'étude	Nombre d'observations
Australie	5	Maths	39	75	1 963	3 mois	8-10
Canada (Ont./angl.)	8	Maths	18	27	757	7 mois	8
Canada (Ont./fr.)	8	Maths	14	18	421	6 mois	8
Canada (Québec)	7-8	Maths	n.d.	30	749	6 mois	10
Hongrie	5	Physique	n.d.	39	1 180	n.d.	5
Israël	6	Histoire	11	22	671	4 mois	1-5
Corée	5	Sciences	15	45	2 400	2 mois	15
Pays-Bas	8	Maths	17	50	1 125	7 mois	8
Nigéria	8	Maths	35	35	1 416	n.d.	4
Thaïlande	5	Maths	77	87	2 572	9 mois	6

Variables et catégories

Le premier défi de taille dans la planification de l'ÉCSC était d'établir une liste des comportements des enseignants qui ferait partie de l'étude descriptive. Les chercheurs participants se mirent d'accord pour n'inclure que les comportements des enseignants répondant aux exigences suivantes : présenter un rapport avec les conclusions touchant l'apprentissage des élèves (tel que démontré par les conclusions des recherches précédentes), relever de la définition et des lignes directrices de l'étude internationale et pouvoir être inculqués aux enseignants dans une période de temps relativement courte. (Ce dernier critère fut choisi en raison de l'intention de procéder à une étude expérimentale une fois l'étude corrélationnelle terminée.) Puisant dans une variété de comportements potentiels identifiés dans plusieurs inventaires de documentation effectués au cours des années soixante-dix (Crawford et Gage, 1977; Dunkin et Biddle, 1974; et Rosenshine, 1976, 1979), les chercheurs de l'AIERS travaillant en collaboration ont retenu des comportements qui semblaient correspondre aux critères. Ensuite, ils les partagèrent en deux catégories, les pratiques de la direction d'une classe (destinées à mettre les élèves en contact avec des tâches d'apprentissage pertinentes) et les pratiques éducatives (destinées à influencer directement les processus d'apprentissage des élèves pendant qu'ils s'adonnent à des tâches d'apprentissage). On choisit aussi des séries de variables sur la collectivité, l'école, les classes, les salles de classe, les élèves, les parents et les enseignants de manière à favoriser une compréhension plus complète du contexte dans lequel l'enseignement allait être observé. Ces décisions firent que l'on inclut quelque 160 variables sur lesquelles allait porter la cueillette de données, à l'aide de plusieurs différents types d'instruments. (En fait, étant donné que chaque réponse à une question, ou réaction à un élément consigné sur une échelle, avait le potentiel de devenir une variable individuelle, il y eut au départ au moins 240 variables discrètes.) Dans l'annexe A, nous donnons une liste des variables ayant effectivement servi dans une ou plusieurs analyses.

Dû au grand nombre de variables à inclure dans l'étude, il fallut établir une forme d'organisation logique des variables, particulièrement aux fins des analyses comportant plusieurs variables qu'il fallait effectuer. Le comité d'orientation de l'étude choisit de classer les variables en quinze groupes ou catégories. Cinq de ces catégories comprenaient des variables démographiques. On les appela «caractéristiques de la collectivité», «caractéristiques de l'école», «traits des enseignants», «caractéristiques du milieu familial» et «caractéristiques des élèves». Cinq autres catégories comprenaient les variables de com-

portements en classe, organisées selon la fonction à laquelle ils répondaient (ou, pour être plus exact, ils devaient répondre) dans la salle de classe. Par exemple, la fonction de certains comportements de l'enseignant ou enseignante ou de certaines activités en classe est de permettre aux élèves d'acquérir des connaissances ou des habiletés. Ces comportements et ces activités furent rangés dans une catégorie nommée «l'enseignement aux élèves». Des regroupements semblables de comportements et d'activités donnèrent des catégories comme «l'orientation des élèves», «la direction des élèves», «l'évaluation des élèves» et «la participation des élèves» (dans ce cas, la participation est du point de vue des élèves et veut dire qu'ils s'engagent activement dans les activités en classe et dans le processus d'apprentissage). Deux catégories concernaient des variables associées aux façons dont les enseignants et les élèves se considéraient mutuellement et percevaient la salle de classe; on les appela «les perceptions des enseignants» et les «perceptions des élèves». Deux autres catégories regroupèrent des variables sur la nature de la classe elle-même et le genre et la quantité d'instruction que les élèves y recevaient réellement; on leur donna le nom de «contexte de la salle de classe» et de «quantité d'apprentissage». Enfin, une catégorie, les «résultats des élèves», touchait les variables des attitudes des élèves et de leur rendement dans la matière enseignée.

Instruments

Un autre gros défi confrontait les participants à l'étude. Il s'agissait de la mise au point d'instruments pour recueillir des données sur toutes les variables mentionnées ci-dessus. En tout, on conçut quatorze instruments à utiliser dans l'ÉCSC; chaque pays participant ne conçut ou ne choisit que deux d'entre eux, le prétest et le post-test des élèves. Les autres instruments s'élaborèrent au niveau international. Tous les pays participants devaient adopter tous les instruments internationaux à l'exception de deux, l'échelle d'évaluation en classe et la formule «Chances d'apprentissage».

Il incombait aux enseignants participants de répondre à cinq instruments. Le questionnaire de départ des enseignants posait 24 questions sur leurs antécédents personnels et leurs responsabilités, l'école, la classe observée, leur autonomie touchant un certain nombre de décisions. Au cours de l'entrevue des enseignants, ceux-ci avaient à lire des descriptions de six approches de direction et de cinq approches d'enseignement, et choisir deux descriptions correspondant le mieux à leurs propres approches de direction et d'enseignement.

Au cours des entrevues précédant l'observation, qui avaient lieu avant chacune

des observations en classe, on demandait aux enseignants de lire plusieurs énoncés et de dire lequel décrivait le mieux leur intention principale ou leur objectif éducatif et l'accent principal qu'ils mettront dans la leçon observée. Le questionnaire final des enseignants comportait 16 questions à réponses multiples sur des points comme la fréquence d'utilisation de divers auxiliaires d'enseignement, l'utilisation de divers tests et activités employés pour l'attribution des notes et le dépistage des difficultés d'apprentissage de même que des questions sur l'utilisation de divers styles et méthodes d'enseignement. La formule «Chances d'apprentissage», facultative, incitait les enseignants à exprimer leurs jugements sur la mesure dans laquelle leurs élèves recevaient un enseignement relié aux connaissances et habiletés évaluées par les différentes questions du post-test des élèves.

Quant aux élèves participant à l'étude, on leur demanda de répondre à quatre instruments. Le questionnaire de départ des élèves consistait en 10 questions sur leurs parents et antécédents personnels et 20 énoncés d'attitude pour lesquels ils devaient indiquer le degré de leur accord ou désaccord. Le prétest, mis au point au niveau national, avait pour but d'évaluer le degré de connaissances et d'habiletés dans la matière examinée, avant l'étude. Le questionnaire final des élèves comportait 36 énoncés, dont certains concernaient le climat de travail dans la salle de classe et d'autres couvraient les mêmes attitudes que celles répertoriées dans le questionnaire de départ des élèves. Encore une fois, les élèves devaient indiquer leur accord ou leur désaccord avec 26 énoncés et répondre vrai ou faux à 10 énoncés. Le post-test, lui aussi conçu au niveau national, cherchait à évaluer la quantité d'instruction acquise par les élèves au cours de la période de l'étude.

Pour terminer, les observateurs en classe devaient remplir quatre instruments pour chacune des classes observées. On se servait de la fiche de renseignements et d'identification de la classe pour consigner l'effectif des classes, le nombre d'élèves présents et les dates de début et de fin de l'enseignement. L'instrument «Instantanés», utilisé cinq fois pendant chaque leçon observée, obligeait l'observateur à noter le type d'activité de classe se produisant (par exemple, cours magistral, révision, discussion, pratique ou exercices oraux, contrôle, lecture individuelle, travail individuel écrit, travaux de laboratoire, activités parascolaires, de direction ou autres), le rôle de l'enseignant ou enseignante pendant l'activité (e.g. interaction, aide, non intervention) et le nombre d'élèves qui semblaient prendre part à l'activité.

L'instrument de mesure de l'interaction sur cinq minutes (ICM), également utili-

sé cinq fois au cours de chaque leçon observée, exigeait que les observateurs codifient et notent les interactions survenues pendant des périodes de cinq minutes. Des codes s'appliquaient à chacun de ces trois aspects : le contexte de l'interaction (e.g. grand groupe, groupe restreint, interaction privée, aide, transition ou non intervention de la part de l'enseignant ou enseignante); la direction de l'interaction (de l'enseignant ou enseignante vers un élève individuel, vers le groupe ou vers un autre adulte ou du groupe vers l'enseignant ou l'enseignante), et la nature même de l'interaction. Dans le cas de ce dernier aspect, les observateurs choisissaient d'abord l'une de ces cinq colonnes : enseignement, question, réponse, réaction, ou direction. À l'intérieur des colonnes, ils indiquaient le code approprié. Ces colonnes permettaient de rapporter 28 catégories ou genres d'interactions (e.g. il donne un exemple, demande une opinion, offre une brève réponse, dit que la réponse proposée n'est pas la bonne). Le quatrième instrument d'observation, le rapport d'activité individuel (RAI), devait aussi être rempli à cinq reprises au cours de chaque leçon observée. À chacune de ces reprises, les observateurs observaient huit élèves choisis au préalable pour savoir combien d'entre eux prenaient part à l'activité ayant lieu à ce moment.

Bref, les observateurs en classe de chaque pays devaient utiliser dans un ordre déterminé trois instruments, les Instantanés, le ICM et le RAI, cinq fois au cours de chacune des leçons observées. Pendant une période de 40 minutes, par exemple, il y aurait cinq séquences de huit minutes chacune. Cinq de ces minutes serviraient à codifier le ICM; les trois minutes suivantes suffisaient à remplir les Instantanés et le RAI. Bien sûr, l'utilisation de ce système d'observation exigeait une formation intensive. Ce système avait été mis au point en grande partie à l'Université Stanford, qui avait utilisé des instruments comparables dans d'autres études «processus-résultats» menées aux États-Unis. Un des chercheurs, Margaret Needels, assura la formation des coordonnateurs d'observation des pays participants lors de séances tenues aux Pays-Bas et en Belgique. Les coordonnateurs des pays dirigèrent ensuite des ateliers de formation semblables, dans leurs propres institutions, pour les personnes qui avaient été choisies comme observateurs potentiels. On s'efforça de recruter des observateurs possédant quelque expérience d'enseignement. Ces ateliers de formation duraient sept jours et comportaient des sessions de pratique en classe de même que des exercices de codification à l'aide de classes filmées sur bandes magnétoscopiques. On ne retint pour effectuer les observations dans le cadre de l'étude que les candidats-observateurs dont les codifications étaient fiables et «valides» (cela était vérifié par des comparaisons avec des codes normatifs).

Pendant l'étude, on procéda aussi de temps à autre à des vérifications pour s'assurer que les mesures relevées répondaient aux critères de fiabilité.

On pensa aussi à un autre instrument de mesure à usage facultatif, l'échelle d'évaluation en classe (ÉÉC), qui comportait des évaluations globales faites après la leçon, sur le caractère approprié d'un ensemble de variables sur la salle de classe (e.g. la clarté de la présentation de l'enseignant ou enseignante). Les résultats relevés sur cette formule ne furent pas intégrés aux analyses internationales.

Le tableau 2 présente un résumé des instruments mis à contribution dans l'ÉCSC de l'AIERS, leur calendrier d'application, ainsi que les personnes qui les ont utilisés. À remarquer que l'annexe A donne une liste des variables soumises aux analyses internationales, par catégorie. Elle donne aussi l'instrument ayant servi à la cueillette de chaque variable.

Préparation et analyse des données

Même si la présente section ne présente pas un intérêt général, elle renseignera sur les tâches monumentales d'organiser, de trier et de publier les données, de présenter les données pour chaque pays et d'effectuer plusieurs sortes d'analyses variées. Cette discussion, nous l'avons voulue aussi brève et simplifiée que possible, mais elle illustre tout de même la complexité de décrire des interactions ponctuelles en classe et de présenter des données amassées par le biais d'études «processus-résultats».

Nous avons déjà mentionné le grand nombre de variables potentielles sur lesquelles des données furent recueillies. On dut en éliminer beaucoup pour des raisons d'ordre technique et conceptuel; on en arrangea d'autres en échelles ou en variables composites. Les catégories organisationnelles servirent aussi à réduire le nombre de variables. Les lecteurs intéressés à connaître les critères et les méthodes employés pour choisir les variables et les organiser en catégories pourront consulter Anderson (sous presse). Malgré tout, d'énormes quantités de données avaient été cueillies dans chacun des pays participants. Prenons par exemple les variables sur l'interaction; il y avait environ 60 codes sur chacun des trois aspects examinés (contexte, direction et nature de l'interaction), pendant chaque période de cinq minutes de leçon. Cette codification se répétait cinq fois par leçon. Dans les cas où l'observation avait porté sur huit leçons, il pouvait y avoir jusqu'à 60 codes X 3 aspects X 5 périodes X 8 leçons = 7 200 codes, pour une seule classe d'un seul pays!

TABLEAU 2

RÉSUMÉ DES INSTRUMENTS UTILISÉS

ÉTUDE SUR LE CONTEXTE DE LA SALLE DE CLASSE DE L'AIERS

QUAND	INSTRUMENT ET ABRÉVIATION	QUI
Avant l'étude	Questionnaire de départ des enseignants (QDE)	Enseignants
	Entrevue des enseignants (EE)	Enseignants
	Questionnaire de départ des élèves (QDÉ)	Élèves
	Prétest (PRÉTEST)	Élèves
Avant chaque observation	Entrevue précédant l'observation (EPO)	Enseignants
Pendant chaque observation	Fiche de renseignements et d'identification de la classe (FRIC)	Observateurs
	Instantanés (INST)	Observateurs
	Interaction sur cinq minutes (ICM)	Observateurs
	Rapport d'activité individuel (RAI)	Observateurs
	Échelle d'évaluation en classe (ÉÉC)*	Observateurs
À la fin de de l'étude	Questionnaire final des enseignants (QFE)	Enseignants
	Chances d'apprentissage (CA)*	Enseignants
	Questionnaire final des élèves (QFÉ)	Élèves
	Post-test (POSTTEST)	Élèves

* Instruments facultatifs.

C'est la raison pour laquelle les chercheurs des pays regroupèrent les données avant de les envoyer au centre de traitement international de l'Université de la Caroline du Sud.

Afin de faciliter l'épuration et la préparation des données, le comité de direction de l'étude convint de cinq règles concernant les données manquantes. Ces règles permettaient de rayer d'un pays les variables, les données sur les élèves, les enseignants et les classes quand les limites sur lesquelles on s'était entendus n'avait pas été respectées. L'une de ces règles, par exemple, disait que les données sur un enseignant ou une enseignante (et donc sur sa classe) seraient éliminées si cet enseignant ou cette enseignante n'avait pas rempli le questionnaire de départ et le questionnaire final. Pour les autres règles, consultez Anderson (sous presse).

Après l'épuration et la préparation des données, on créa trois groupes de données pour chacun des pays : un groupe de données au niveau des élèves, un au niveau des classes et un au niveau des pays. Lors de la création du groupe de données sur les élèves, on assigna bien sûr à chaque élève sa propre réponse (sous forme de résultat ou code numérique) sur les variables des quatre instruments «élèves». Pour chacune des variables des quatre instruments «élèves», on assigna à chaque élève une valeur numérique correspondant à la réponse de son enseignant. Pour chacune des variables des quatre instruments d'observation, on attribua à chaque élève la valeur moyenne (ou la statistique appropriée) de la classe, sur chaque période d'observation, et répartie sur toutes les périodes d'observation.

Le groupe de données au niveau des classes était le même que celui au niveau des élèves sauf dans le traitement des variables des instruments «élèves». Pour ces variables, on utilisa la réponse ou le résultat moyen (ou la statistique appropriée) de tous les élèves de chaque classe.

Enfin, de façon à déterminer les similitudes et les différences entre les pays participants, on créa des groupes de données au niveau des pays en rassemblant les données relatives à tous les élèves, enseignants, classes de chaque pays. Ces groupes de données comprenaient des statistiques à une seule variable comme les écarts moyens et standard dans le cas de certaines variables et dans le cas d'autres variables, les fréquences et les pourcentages, tel qu'approprié. Le travail que nous venons de décrire est l'oeuvre de Garrett Mandeville, de l'Université de la Caroline du Sud. Une fois qu'il eut fini son travail, il envoya ses groupes de données pour chaque pays à Norbert Sellin, de l'Université de

Hambourg où l'on se livra à des analyses basées sur plusieurs variables.

En préparant les données pour les analyses basées sur plusieurs variables, on examina attentivement les données brutes. On prépara des caractéristiques de dispersion et des corrélations des résultats aux prétests avec ceux des post-tests, pour chaque classe, de façon à identifier les classes dont les résultats aux prétest et post-test présentaient des relations particulières. On prit note des classes qui avaient obtenu des résultats moyens au post-test inférieurs à ceux du prétest d'au moins quatre points. On identifia les classes qui avaient des taux élevés de non-réponse à des instruments particuliers. On examina les distributions de fréquence de toutes les variables et on prit note des variables dont les distributions étaient fortement faussées. On identifia les variables ayant peu ou pas du tout de variance. On prit en considération ces particularités au moment de l'interprétation des corrélations entre les variables d'observation et de rendement des élèves. En se fondant sur cet examen attentif et sur les discussions entre les participants sur les résultats de cet examen, on élimina des classes ou des variables des analyses plus poussées au niveau des pays. Cependant, on inclut les résultats de cet examen dans les données descriptives relatives à chaque pays.

Ce travail préliminaire achevé, on se livra à deux analyses à plusieurs variables. Ces analyses examinèrent la nature et la vigueur des rapports entre les variables choisies. Sans entrer dans le détail (comme le font Anderson, sous presse, et le rapport international, à paraître), nous allons décrire ces deux analyses brièvement.

La première analyse à plusieurs variables utilisait les groupes de données au niveau des élèves et comprenait 15 variables prises principalement dans les instruments «élèves». Donc, on l'appela le modèle de variables «élèves». Voici les 15 variables : antécédents familiaux, genre d'école, âge de l'élève, sexe de l'élève, prétest, aspirations de l'élève vers une éducation de niveau supérieur, attitude envers l'école, prédisposition pour la matière, temps consacré par l'élève, perception par l'élève de l'orientation de la classe vers un but, perception par l'élève de l'efficacité du cadre d'apprentissage créé par l'enseignant ou l'enseignante, perception par l'élève du degré d'appréciation du travail et des progrès accomplis exprimé par l'enseignant ou l'enseignante, chances d'apprentissage (qu'on ne fournit que pour les pays ayant choisi d'utiliser l'instrument CA), et les deux variables de résultats : le post-test de l'élève et son attitude envers la matière, après l'étude. On testa le modèle de variables «élèves» dans six des pays participants : l'Australie, le Canada, la Hon-

grie, la République de Corée, les Pays-Bas et la Thaïlande. Pour le Canada, les deux versions ontariennes (anglaise et française) en faisaient partie; il y avait donc sept versions. On exclut les autres pays (y compris le Québec), en raison d'échantillons trop petits, de données insuffisantes, ou du manquement apparent de fiabilité des variables clés comprises dans le modèle.

On testa le modèle mis au point pour la deuxième analyse à plusieurs variables à la fois dans les groupes de données aux niveaux des élèves et des classes. On l'appela le «modèle noyau» de l'étude. Il s'agissait essentiellement d'une version élargie du modèle de variables «élèves». Aux 15 variables «élèves», on ajouta six autres variables empruntées aux instruments des enseignants et des observateurs. Voici ces six variables : taille de la classe, années d'expérience dans l'enseignement, perception par l'enseignant ou l'enseignante des élèves dans la classe (e.g. niveau d'habileté, pourcentage d'élèves ayant besoin d'enseignement correctif), enseignement direct (e.g. donner des indices, des explications, des exemples), pratiques d'interrogation (e.g. poser des questions, commenter les réponses des élèves), et la direction de la classe (e.g. s'engager dans des activités méthodiques, comme le ramassage des devoirs, le fait de discipliner les élèves se conduisant mal). Ces variables supplémentaires furent choisies en partie en raison des analyses préliminaires de leurs corrélations simples et partielles avec le rendement et les attitudes des élèves. Dans ce modèle donc, le groupe de variables sur les enseignants et les classes représentait l'estimation empirique la plus juste possible des variables du contexte de la salle de classe pouvant influencer le rendement et les attitudes des élèves. Malheureusement, le «modèle noyau» ne put être testé que dans quatre pays étant donné qu'il exigeait au moins 40 classes (après l'épuration et la préparation des données). Ces quatre pays étaient l'Australie, la République de Corée, les Pays-Bas et la Thaïlande.

Dans les deux modèles, on décrit les liaisons entre les variables sous la forme d'un modèle causal ou d'un schéma de chemin critique. On utilisa la méthode statistique des moindres carrés partiels pour l'analyse des données selon ces modèles. Voir Anderson (sous presse) ou le rapport international à paraître pour une discussion sur la méthode des moindres carrés partiels et pour des détails sur les premières analyses, les réanalyses et les changements apportés aux modèles pour concevoir des modèles «propres aux pays». Cependant, il convient de noter ici que «l'analyse du chemin critique que l'on a appliquée à ces modèles peut être vue simplement comme l'amélioration des analyses utilisées dans les études processus-résultats antérieures dans lesquelles les caractéristiques fixes sont partagées à partir des rapports entre les mesures des comportements

des enseignants et le rendement noté au post-test» (Mandeville, 1984, p. 156).

Principales conclusions et généralisations tirées de l'étude sur le contexte de la salle de classe (ÉCSC)

Malgré le soin apporté à la préparation des données et au choix de techniques d'analyse appropriées de ces données, les résultats de l'étude de l'AIERS furent quelque peu décevants. Du point de vue descriptif, les données obtenues grâce aux instruments d'observation, particulièrement l'instrument interaction sur cinq minutes, ne parvinrent pas à différencier les pratiques éducatives en vigueur dans les différents pays. Par exemple, des 28 catégories d'interactions du ICM, six rendaient compte de 66 % à 77 % de toutes les interactions observées dans presque tous les pays participants. Presque la moitié des catégories d'interactions ne furent que rarement, sinon jamais, utilisées par les observateurs dans quelque pays que ce soit. Dans le cadre de cette étude qui devait comporter une phase corrélationnelle et une phase expérimentale, ces 13 comportements d'interactions (e.g. pose une question d'ordre supérieur, vérifie l'efficacité d'une explication), n'auraient pu faire partie de la phase expérimentale, car ils n'auraient pu corréler avec le rendement des élèves, à cause de leur faible fréquence et de leur manque de variance.

Du point de vue inférentiel, les données recueillies par les instruments d'observation n'ont pas produit d'effets marqués et directs en termes de rendement des élèves dans le contexte du modèle «noyau». Dans quelque pays que ce soit, on n'observa que rarement des coefficients du chemin critique supérieurs à 0,10. Même les corrélations d'ordre zéro des variables d'observation avec le rendement étaient faibles. De plus, beaucoup des corrélations les plus fortes pouvaient s'attribuer principalement à des irrégularités dans la distribution des données sur des variables d'observation particulières (e.g. la présence d'un unique enseignant «témoin», variables fortement faussées).

Dans un livre à paraître édité par Ryan, Anderson et Shapiro, nous disposerons d'une présentation complète et de discussions des conclusions de l'étude sur le contexte de la salle de classe. Le bref résumé de quelques-unes des conclusions que nous offrons ci-dessus avait pour but de préparer le lecteur aux six généralisations que nous avons dérivées de nos conclusions. Deux de ces généralisations proviennent des analyses préliminaires des données, l'une est dérivée des statistiques à une seule variable. Les trois autres généralisations sont tirées des analyses statistiques à plusieurs variables réalisées dans le cadre du modèle de variables «élèves» et du modèle «noyau». Nous citons directement Anderson

(sous presse) pour la présentation et la discussion de ces généralisations.

Généralisation 1. À l'intérieur des pays, il y a d'importantes différences entre ce que les enseignants enseignent et ce qui fait l'objet des tests.

Comme choix au niveau des pays, on demanda aux enseignants d'indiquer à quel point leurs élèves avaient eu la chance d'apprendre la matière contenue dans chaque question du post-test. En se fondant sur leurs réponses, on élabora pour chaque enseignant ou enseignante un index des chances d'apprentissage (CA). Cet index représentait le pourcentage de questions du post-test des élèves que, selon l'enseignant ou l'enseignante, les élèves avaient eu la chance d'apprendre pendant la durée de l'étude. Par exemple, un index CA de 75 indiquait qu'un enseignant ou une enseignante croyait que ses élèves avaient eu la chance d'apprendre 75 % des questions du post-test.

Comme le choix de l'instrument CA avait été laissé aux pays, on ne disposait pas de données CA pour tous les pays. En conséquence, Mandeville (1984) conçut une méthode empirique d'évaluation des CA. Le raisonnement de Mandeville était le suivant : si les élèves avaient eu la chance d'apprendre le contenu testé par une question particulière du test, alors au moins cinq autres pour cent des élèves devraient être en mesure de répondre correctement à une question du post-test, par rapport au prétest. Aussi, on pourrait calculer le pourcentage des questions pour lesquelles l'augmentation dans la proportion du nombre d'élèves répondant correctement à chaque question dépassait cinq pour cent. Ce pourcentage de questions représenterait une estimation empirique des CA.

Les deux estimations des CA comportent d'évidentes sources d'erreurs. Les estimations empiriques des CA ne sont pas indépendantes des résultats aux post-tests et sont soumises à l'influence de tout facteur ayant pu influencer l'apprentissage par les élèves d'éléments particuliers (par exemple, les élèves particulièrement doués capables de déduire les réponses à des questions sur de la matière non couverte en classe, les enseignants exceptionnels qui aident les élèves à généraliser leurs connaissances ou à les appliquer à des situations nouvelles et différentes). D'un autre côté, les estimations des CA fondées sur des jugements ne sont pas étrangères aux sentiments des enseignants lorsqu'ils remplissent l'instrument CA. En particulier, les enseignants qui se sentent évalués pourront répondre que, tel qu'attendu, ils enseignent à leurs élèves tout ce qui a été testé.

Alors, à la lumière de ces sources d'erreurs, on pourra trouver que l'estimation

empirique (fixée à une augmentation d'un élève sur vingt) sous-estime les CA réelles. On pourra trouver que l'estimation fondée sur des jugements est surestimée (en raison de la pression exercée sur les enseignants pour qu'ils donnent des réponses socialement acceptables). Comme c'est fréquemment le cas donc, l'estimation générale la plus «exacte» possible pourrait se situer à mi-chemin entre ces deux estimations. En fait, les données relatives à ces deux estimations alimentent le débat que nous venons d'exposer. En général, l'estimation empirique des CA est approximativement de 10 à 15 % inférieure à l'estimation fondée sur des jugements (dans les pays pour lesquels on dispose d'estimations empiriques et fondées sur des jugements).

Cependant, indépendamment de l'estimation, la variation des chances d'apprentissage en classe à l'intérieur des pays est frappante. En utilisant l'estimation par jugements, l'index des CA dans quatre des pays participants varie de 0 à 100, de 8 à 100, de 28 à 97, et de 30 à 100.[1] (Les moyennes dans ces quatre pays sont respectivement de 79, 67, 75 et 87.) En utilisant les estimations empiriques, l'index des CA dans quatre des pays participants varie de 20 à 88, de 29 à 97, de 40 à 93 et de 48 à 100, pour des moyennes respectives de 52, 63, 65 et 85.)

Quel que soit le pays, et quelle que soit la méthode d'estimation employée, la différence entre les classes à index CA élevé et celles à index CA faible n'est jamais inférieure à 50 %; c'est dire que dans certaines classes, les élèves ont eu la chance d'apprendre de deux à trois fois plus d'éléments du test que les élèves dans d'autres classes.

Alors, dans presque tous les pays participants, les élèves de classes différentes révèlent d'importantes différences dans les chances qu'ils ont eues au cours du post-test de démontrer la matière qu'ils avaient acquise.

L'implication de cette généralisation est bien évidente. Dans la mesure où les chances d'apprendre influencent le rendement des élèves, les élèves placés dans

---

[1]Les lecteurs trouveront peut-être intéressantes les conclusions de la version ontarienne réalisée dans les écoles anglophones. Pour notre échantillon d'enseignants, l'index des CA fondé sur leurs propres rapports variait de 28 % à 97 %. Ce qui veut dire qu'un enseignant ou une enseignante a rapporté que seulement 28 % des éléments du test correspondaient à ce qu'il ou elle avait enseigné au cours de l'année, tandis que pour l'autre enseignant ou enseignante, les éléments du test correspondaient de près à ses objectifs d'enseignement. La moyenne fut de 75 % et la médiane de 80 %. Les éléments du test provenaient du test de l'AIERS pour les mathématiques et on les avait trouvés valides pour le programme de mathématiques en 8e année de l'Ontario.

des classes différentes vont obtenir des résultats différents dans des tests visant à mesurer leur rendement, indépendamment du genre ou de la qualité de l'enseignement qu'on leur donne. Par conséquent, l'index CA devient une variable importante qui n'entre pas en ligne de compte dans notre quête de pratiques et de comportements éducatifs efficaces.

Généralisation 2. Les enseignants se conduisent avec inconsistance d'un jour à l'autre.

Lorsque nous essayons de différencier les enseignants en termes de pratiques, de stratégies et de styles, ce serait une bonne chose si nous pouvions compter sur les enseignants pour n'avoir que des comportements constants au fil des jours, des leçons ou des périodes d'observation. Cette constance permettrait de classer de façon fiable les enseignants selon des typologies diverses. Les résultats de l'ÉCSC indiquent que ce n'est pas le cas. On a trouvé que les enseignants manquaient tout à fait de suite dans leurs comportements d'un jour à l'autre. Disons-le simplement, les enseignants disent et font des choses différentes d'un jour d'observation à l'autre. De plus, ils organisent des activités différentes (e.g. cours magistral, travail individuel, exercices oraux) d'une observation à l'autre.

Cette généralisation est étayée par une statistique appelée «carré êta», qui représente la proportion de variance dans les comportements des enseignants, pouvant être attribuée à des différences entre les enseignants eux-mêmes. Plus le coefficient de carré êta s'amenuise, plus il devient difficile de différencier de façon fiable les enseignants en termes de comportements, et plus il est vraisemblable que les différences de comportement puissent être attribuées à des différences dans les arrangements, les situations, les jours de la semaine et d'autres facteurs semblables. Quand le carré êta est inférieur à 0,50, la variance dans les comportements de l'enseignant ou enseignante attribuable aux différences entre les enseignants équivaut à la variance attribuable à tous les autres facteurs ou sources. Quand le carré êta est inférieur à 0,50, la variance dans le comportement de l'enseignant ou enseignante attribuable à d'autres facteurs (e.g. les arrangements, situations, jours) est supérieure à la variance attribuable aux différences entre les enseignants eux-mêmes.

En gardant ces questions à l'esprit, considérons les conclusions de l'ÉCSC Penchons-nous, par exemple, sur les résultats relatifs à une seule activité, les cours magistraux. Dans aucun pays, le carré êta pour les cours magistraux ne dépassa 0,50; c'est dire que dans aucun pays, la tendance à faire des cours ma-

gistraux ne peut être attribuée principalement aux différences entre les enseignants. La tendance à faire des cours magistraux pourrait plutôt s'attribuer surtout à d'autres facteurs (comme le jour de l'observation -- peut-être que les enseignants avaient tendance à faire plus de cours magistraux le lundi que le vendredi; les résultats recherchés par la leçon observée -- peut-être les enseignants étaient-ils plus portés à faire des cours magistraux lorsqu'ils enseignaient des faits que lorsqu'ils montraient comment résoudre des problèmes; et l'organisation des salles de classe -- peut-être les enseignants avaient plus tendance à donner des cours magistraux lorsqu'ils disposaient leurs classes en rangées que lorsqu'ils les disposaient en cercle, pour les discussions de groupes). On trouva ce même résultat (c'est à dire des coefficients de carré êta inférieurs à 0,50) dans le cas de 11 des 52 variables (approximativement 21 %) comprises dans les instruments d'observation.

Par rapport aux autres variables comprises dans les instruments d'observation, le coefficient dépassait 0,50 dans un seul des pays participants pour 18 variables (ou approximativement 35 %). Ainsi, pour un peu plus de la moitié des variables comprises dans les instruments d'observation, la variabilité des comportements des enseignants au cours des périodes d'observation était observable dans tous les pays, ou dans tous les pays sauf un. En fait, pour une seule variable, notamment la façon dont l'enseignant ou l'enseignante reconnaît la correction des réponses des élèves aux questions, peut-on dire qu'il existait des différences assez fiables entre les enseignants dans la majorité des pays participants (cinq pays avaient des coefficients de carré êta supérieurs à 0,50).

La conclusion que les différences observées dans les comportements des enseignants ne peut pas, de façon fiable, être attribuée aux différences entre les enseignants, jette un doute sur la possibilité d'utiliser les résultats de l'ÉCSC pour concevoir des études expérimentales simples afin d'examiner les effets des pratiques ou des styles des enseignants sur le rendement des élèves et les attitudes des enseignants. Les résultats suggèrent plutôt qu'une série de variables contextuelles (e.g. les différences dans les objectifs éducatifs, les types de leçons enseignées, l'organisation des classes, les jours de la semaine) contribuent davantage aux différences dans les comportements des enseignants que les différences caractéristiques entre les enseignants (e.g. un ensemble de comportements préférés). En fait, les résultats semblent indiquer que les enseignants adaptent vraisemblablement leurs comportements aux variables situationnelles pertinentes de façon à ce qu'ils soient <u>appropriés</u> plutôt que <u>cohérents</u>. Ainsi, les programmes de formation des enseignants devraient, si l'on veut qu'ils soient couronnés de succès, avoir deux buts : (1) aider les enseignants à

développer un répertoire d'aptitudes et de comportements pédagogiques qui pourront leur être utiles, et (2) aider les enseignants à développer la capacité de choisir le moment d'adopter telle aptitude ou tel comportement en classe.

Généralisation 3. La nature de l'enseignement en classe se ressemble d'un pays à un autre.

Au premier abord, cette généralisation semble contredire la précédente. Si des enseignants diffèrent si grandement dans leurs comportements en diverses circonstances, comment est-il possible de parler de ressemblance de l'enseignement? On peut résoudre ce paradoxe apparent de deux manières. Premièrement, cette généralisation est fondée sur des données rassemblées au niveau des pays, c'est à dire qu'elle se fonde sur le pourcentage moyen de temps passé en activités diverses ou consacré à toutes les interactions de comportements de tous les enseignants au cours de toutes les périodes d'observation. Deuxièmement, même si les enseignants ont tendance à être inconsistants dans les façons dont ils se comportent au cours d'une longue période de temps, le répertoire de comportements qu'ils manifestent dans leurs classes, quelle qu'en soit la fréquence, est relativement restreint. L'incohérence observée est causée par les différences dans leur usage d'un nombre limité d'activités et de comportements.

Pour illustrer ce dernier point, prenons l'exemple suivant. L'instrument INST comprenait une variable appelée activités. Dix types ou catégories étaient comprises dans l'instrument (e.g. cours magistral, discours, exercices oraux, travail individuel). En général, les résultats suggéraient que les trois principaux types d'activités survenaient à un degré plus ou moins élevé dans les classes de tous les pays participants. Les enseignants parlent «aux» élèves ou «avec» leurs élèves (de 35 % à 82 % du temps selon les pays); les élèves travaillent à leur bureau ou table de laboratoire (de 15 % à 47 % du temps); et les enseignants se livrent à toute une gamme d'activités générales reliées à la direction de la classe, comme prendre les présences, distribuer ou ramasser des devoirs (de 1 % à 16 % du temps).

Les différences entre les pays avaient tendance à se retrouver dans ces trois catégories générales d'activités. Dans certains pays, les enseignants étaient portés à parler «aux» élèves (i.e. cours magistral); dans d'autres, ils parlaient «avec» les élèves (i.e. discussion). Dans certains pays, les discussions étaient davantage centrées sur de la nouvelle matière; dans d'autres, les discussions faisaient le point sur la matière déjà vue. Dans la plupart des pays, les élèves passaient beaucoup de temps à leur bureau ou table de laboratoire, à

faire le travail qui leur avait été assigné. Dans la plupart des pays, ces travaux étaient écrits, bien que dans deux des pays, ceux qui avaient choisi les sciences pour l'étude, les travaux étaient en partie des projets ou des expériences.

Des résultats étayant cette généralisation proviennent aussi des données sur la nature des interactions entre les enseignants et les élèves colligées par l'instrument ICM. Ce dernier comprenait 28 catégories d'interactions (e.g. l'enseignant ou l'enseignante utilise un exemple, examine, pose une question d'ordre supérieur, donne une réponse complète à un élève, ne fait pas d'interactions verbales, sociales ou organisationnelles). Sauf pour un pays, six de ces catégories d'interactions rendaient compte pour plus de la moitié des interactions en salle de classe (dans une proportion variant de 56 % à 77 % selon le pays. Mises ensemble, ces six catégories confirment les conclusions sur les activités que nous avons déjà présentées.

Deux de ces catégories ont trait aux explications données par les enseignants et deux aux questions des élèves, aux réponses des enseignants (principalement des questions sur des faits). Ces quatre catégories d'interactions sont reliées à l'activité générale des enseignants de parler «aux» élèves ou «avec» les élèves. L'une des six catégories d'interactions fréquemment codifiées traitait de l'absence d'interactions verbales, une situation qui se produit fréquemment lorsque les élèves font du travail. La dernière catégorie d'interactions fréquemment codifiée fut baptisée «interactions organisationnelles»; il s'agissait du type d'interactions ayant lieu lorsque l'enseignant ou l'enseignante dirigeait les «affaires» de sa classe, ou, pour reprendre l'expression déjà employée, qu'il ou elle s'occupait de la direction de sa classe. Bien qu'il y ait eu encore une fois certaines différences dans la fréquence d'utilisation de ces différents types ou catégories d'interactions dans les divers pays, les ressemblances entre les pays étaient frappantes.

Une autre façon d'exprimer cette troisième généralisation est de dire que la nature de l'enseignement en classe dans les différents pays présente plus de ressemblances que de différences. On ne s'attendait guère à cette généralisation face aux grandes différences sociales et culturelles existant entre les pays participants.

Généralisation 4. Le passé de l'apprenant exerce une forte influence sur son rendement et ses attitudes.

Dit simplement, l'apprentissage précédent et les attitudes avec lesquelles les élèves se présentent à des classes particulières influencent ce qu'ils apprennent dans ces classes et dans les cours, et comment ils se sentent par rapport à la matière enseignée. En général, cette influence est tellement forte que peu de variables peuvent compenser ses effets, et encore moins les surmonter.

Bien que cette généralisation ne soit pas nouvelle dans le monde de l'éducation, les résultats de l'ÉCSC contribuent à nous la faire mieux comprendre de deux façons. Premièrement, le passé de l'apprenant est beaucoup plus complexe que les connaissances, habiletés ou attitudes particulières qui sont mises à profit dans l'étude d'une matière donnée à un moment précis. Le genre de foyer dans lequel l'enfant est élevé, les aspirations qu'entretiennent les élèves sur leur éducation future et traditionnelle et les attitudes qu'ils développent envers l'école en général ont tous tendance à influencer le rendement et les attitudes des élèves dans des matières précises dans presque tous les pays. En général, les résultats donnent à penser que l'apprentissage ou le manque d'apprentissage sont cumulatifs. Bien que la forte influence des antécédents sur les connaissances et les sentiments actuels soit quelque peu diminuée par l'utilisation de mesures différentes au début et à la fin d'une période d'études (tel que ce fut le cas dans quelques pays), l'influence qu'ont les années d'apprentissage ou d'absence d'apprentissage ne sera vraisemblablement pas renversée au cours d'une période relativement courte, quelles que soient l'intensité ou la puissance des expériences éducatives qui seront vécues.

Deuxièmement, l'influence des antécédents familiaux porte principalement sur le rendement et attitudes constatés chez les élèves au début de la période d'enseignement, et non pas sur le rendement et les attitudes des élèves lorsqu'ils quittent la classe ou terminent le cours. Autrement dit, le contexte familial ne semble pas influencer directement l'apprentissage des élèves sur une période particulière. Le milieu familial semble plutôt avoir un rôle de préparation, non d'actualisation. Par conséquent, les effets du milieu familial sur le rendement et les attitudes des élèves peuvent être minimisés si on arrive à trouver des façons d'intervenir avec force entre le début et la fin d'une période d'études.

Généralisation 5. La perception qu'ont les élèves de leur classe et de l'enseignement qu'ils y reçoivent influence leur rendement et leurs attitudes.

Dans tous les pays, la perception qu'ont les élèves de l'objectif de leur classe, c'est à dire à quel point ils perçoivent que leur classe est animée d'un

but éducatif sérieux, exerce une influence faible mais constante sur leur rendement et leurs attitudes. De plus, la perception qu'ont les élèves de la structure d'apprentissage créée par leurs enseignants exerce également une influence faible mais soutenue sur leur rendement et leurs attitudes. La création d'une structure d'apprentissage signifie que l'enseignant ou l'enseignante met l'accent sur les idées importantes et les points principaux, qu'il ou elle fournit des explications claires, incite les élèves à réfléchir pour trouver la solution à des problèmes et qu'il ou elle fait périodiquement des résumés des principales questions étudiées. L'influence de ces perceptions est en quelque sorte plus grande et plus constante dans le cas des attitudes des élèves que dans le cas de leur rendement.

On peut tirer au moins deux conclusions de cette généralisation. Premièrement, les perceptions des élèves de ce qui se passe en classe sont peut-être plus importantes que ce qui s'y produit réellement. La dernière généralisation va d'ailleurs donner plus de poids à cette conclusion. Deuxièmement, dans la recherche de meilleures conditions d'apprentissage dans les salles de classe et de pratiques d'enseignement plus efficaces, il semble logique de commencer par la création de classes dans lesquelles on a mis de l'avant une norme sociale favorisant le travail scolaire et un apprentissage de qualité, et dans lesquelles les enseignants établissent la structure nécessaire pour que les élèves puissent séparer le bon grain de l'ivraie.

Généralisation 6. Les comportements des enseignants et les activités observés en classe semblent n'avoir pratiquement aucune influence sur le rendement et les attitudes des élèves.

Comme nous l'avons déjà dit, on mit au point deux modèles pour la préparation des analyses à plusieurs variables. Le premier comprenait principalement des variables sur les élèves (e.g. antécédents familiaux, attitudes de départ envers l'école, rendement précédent dans la matière enseignée). En plus de ces variables «élèves», le deuxième modèle comprenait un groupe de variables associées aux questionnaires des enseignants et aux instruments d'observation. On est parvenu à une estimation de l'influence des comportements des enseignants et des activités observées en classe sur le rendement et les attitudes des élèves en comparant la proportion de variance dans le rendement des élèves, dont rendait compte le deuxième modèle, avec la proportion de variance dans le rendement des élèves dont rendait compte le premier modèle. Cette comparaison fut possible pour les quatre pays dont on examina les deux modèles (l'Australie, la République de Corée, les Pays-Bas et la Thaïlande).

Dans ces quatre pays, le modèle de variables «élèves» rendit compte de 34 % à 56 % de la variance dans le rendement des élèves, et de 22 % à 42 % de la variance dans les attitudes des élèves. Quand on ajouta les variables de la taille de la classe, des années d'expérience, des perceptions des enseignants sur les habiletés de leurs élèves, de l'enseignement direct, de l'interrogation, et de la direction de la classe au modèle de variables «élèves» (pour créer le modèle noyau), l'augmentation en pourcentage de la variance rendue compte par ces six variables supplémentaires allait de 0 % à 2 % dans le cas du rendement des élèves et de 0 % à 6 % dans le cas de leurs attitudes.

Cette généralisation se renforce davantage par les simples corrélations d'ordre zéro, c'est à dire la liaison entre deux variables. Pour l'ensemble de comportements compris dans la variable «enseignement direct», les corrélations d'ordre zéro avec le rendement des élèves pour les quatre pays étaient de -0,03; 0,02; 0,07 et 0,21. Pour l'ensemble de comportements compris dans la variable «interrogation», les corrélations d'ordre zéro avec le rendement des élèves pour les quatre pays étaient de -0,05; 0,01; 0,07 et 0,10. Enfin, pour l'ensemble de comportements compris dans la variable «direction de la classe», les corrélations d'ordre zéro avec le rendement des élèves pour les quatre pays étaient de -0,06; -0,13; -0,17 et -0,22. Ce dernier groupe de corrélations, qui sont les plus fortes et les plus constantes, suggèrent que le rendement des élèves est moindre dans les classes où l'enseignant ou l'enseignante consacre plus de temps à des comportements relevant de la direction de la classe.

Cette dernière généralisation, bien que troublante à première vue, compte tenu du deuxième but principal de l'ÉCSC, ne causa pas une bien grande surprise, lorsque mise en relief par plusieurs des autres généralisations. Les élèves de classes différentes présentaient d'importantes différences dans leurs chances d'apprendre la matière contenue dans le dernier test mesurant leur rendement. En conséquence, ce que les enseignants et les élèves disaient ou faisaient en classe pouvait ou non être relié à la matière que les élèves devaient apprendre de manière à faire bonne figure au test de rendement. De plus, les enseignants différaient grandement dans leur utilisation d'activités et de comportements particuliers d'une observation à la suivante. Donc, la réunion de données d'observation recueillies à différents moments introduisait vraisemblablement dans les données une bonne quantité d'erreurs. Ces erreurs réduisaient finalement la possibilité de trouver d'importantes corrélations entre les activités et les comportements d'une part, et le rendement et les attitudes des élèves d'autre part.

Si l'on se fie aux données rassemblées, les enseignants étaient assez semblables quant au pourcentage de temps qu'ils consacraient à des activités particulières ou pendant lequel ils démontraient des comportements particuliers. De façon plus importante, il y avait des comportements qu'on voyait rarement chez la grande majorité des enseignants de tous les pays (e.g. utiliser des exemples, poser des questions auxquelles les élèves doivent répondre en exprimant leur opinion, dire que l'affirmation ou la réponse d'un élève sont incorrectes, répondre aux questions des élèves, vérifier si les élèves ont bien saisi une explication, souligner l'importance de ce dont la classe discute). Ces variables avaient souvent des pourcentages moyens proches de zéro et il y avait très peu de variance entre les enseignants. En conséquence, la possibilité que ces variables corrèlent avec le rendement et les attitudes des élèves n'existait pratiquement pas.

Deux analyses secondaires des données de l'ÉCSC

Divers participants à l'étude entreprirent des analyses supplémentaires des données de l'ÉCSC. Je vais décrire brièvement deux d'entre elles. La première fut menée sur les données des classes de mathématiques de la 8e année comprises dans l'échantillon ontarien des écoles anglophones.

Hanna et Ryan (1985) identifièrent d'abord des classes où le progrès scolaire pendant l'année scolaire avait été nettement plus élevé que dans les autres classes composées d'élèves semblables. Pour l'échantillon total de 27 classes, les moyennes du prétest variaient de 13,5 à 19,8 dans un test de 40 éléments; les moyennes du post-test variaient de 17,3 à 29,2; le progrès scolaire moyen pendant l'année scolaire variait de 2,8 à 15,6. Les élèves dans la classe de l'enseignant ou enseignante 331 avaient obtenu une moyenne de 13,6 au prétest, la deuxième plus faible de toutes les classes de notre échantillon. Leur moyenne au post-test était de 29,2, la meilleure de toutes les classes. Leur progrès scolaire moyen était de 15,6, et presque tous les élèves de cette classe avaient réalisé d'importants progrès scolaires pendant l'année. Les élèves de l'enseignant ou enseignante 101 avaient des moyennes de 16,7 au prétest et de 26,6 au post-test, une moyenne de 9,9, la deuxième plus élevée de notre échantillon. Ces deux enseignants (331 et 101) étaient nos enseignants «à progrès scolaire élevé».

En plaçant les enseignants 331 et 101 avec les autres enseignants dont les élèves avaient obtenu des résultats moyens similaires au prétest, nous avons comparé les enseignants «à progrès scolaire faible» avec les enseignants «à progrès

scolaire élevé» en ce qui a trait aux diverses caractéristiques des élèves, des classes et des enseignants. Nos analyses indiquèrent que les progrès significatifs en mathématiques dans les classes «à progrès élevé» ne pouvaient pas être attribués aux antécédents des élèves ou des enseignants. Par exemple, les élèves de la classe 331 venaient de familles de faible statut socio-économique et, selon l'enseignant ou l'enseignante 331, presque tous étaient passés en 8e en ayant besoin d'enseignement correctif. En plus des mathématiques, l'enseignant ou l'enseignante 331 enseignait quatre autres matières et n'était pas un(e) spécialiste en mathématiques. À aucun égard, les élèves des classes 331 et 101 ou les deux enseignants ne différaient de façon significative des élèves et des enseignants des autres classes, ce qui aurait pu expliquer leur progrès scolaire élevé.

Cependant, les analyses révélèrent de nombreuses différences entre les stratégies d'enseignement des enseignants «à progrès scolaire élevé» et celles de leurs collègues. Ces différences ne furent pas facilement visibles dans les données produites par l'instrument de mesure de l'interaction sur cinq minutes, lorsque les données de toutes les occasions et leçons étaient réunies. Ces différences apparaissaient dans les intentions éducatives poursuivies pendant les leçons observées, telles qu'exprimées par l'enseignant ou l'enseignante au cours d'entrevues avant l'observation, dans les formes d'activités et les rôles assumés par l'enseignant ou l'enseignante que l'on a pu observer en classe (grâce à l'instrument «Instantanés»), de même que dans les rapports des enseignants sur leurs stratégies et pratiques «préférées» (qu'ils avaient pu exprimer grâce au questionnaire final et à l'entrevue). Nous avons trouvé que les enseignants «à progrès scolaire élevé» adoptent, c'est du moins le cas de l'enseignant ou enseignante 331, une approche de l'enseignement extrêmement organisée, qui signifie que la matière est enseignée jusqu'à ce que les élèves la maîtrisent, ce qui réduit le besoin de réviser fréquemment le travail. Dans le cas de l'enseignant ou enseignante 101, nous avons constaté une approche qui mettait l'accent sur la présentation de la matière suivie d'exercices approfondis permettant d'appliquer la matière vue à des situations nouvelles. Les deux enseignants passaient très peu de temps à la direction de la classe. Les deux travaillaient activement avec les élèves faisant des exercices individuels, qui servaient aussi à vérifier de façon informelle si les élèves avaient compris la matière. Voyons ce que disait l'enseignant ou l'enseignante 331 pendant son entrevue :

> Mes objectifs sont à court terme. Je vérifie toujours si les élèves ont bien compris la matière que nous venons de couvrir. Même si j'enseigne à l'ensemble de la classe, je m'occupe beaucoup des élèves de

façon individuelle. Je vais et je viens dans les allées pendant qu'ils travaillent pour qu'ils ne soient pas frustrés s'il y a quelque chose qu'ils n'ont pas compris. Cela prévient les problèmes de discipline... Jamais je ne répète une unité, et je ne fais pas de contrôle sur une unité avant que les élèves sachent bien la matière qu'elle contient. Lorsque je m'aperçois qu'ils n'ont pas bien compris, j'utilise une autre méthode. J'utilise beaucoup de stratégies diverses. (traduction libre)

Comparez cette approche avec celle de l'enseignant ou enseignante 101, qui est différente, mais qui a semblé efficace avec son groupe particulier d'élèves :

Après que j'aie présenté la matière et donné quelques exemples, je leur demande de faire du travail individuel. L'enseignement de la matière constitue une leçon qui peut occuper toute une période de cours, puis ensuite, je leur laisse une certaine période de temps pour faire les exercices. Règle générale, je prépare à leur intention un grand nombre d'exemples sur lesquels ils peuvent travailler. (traduction libre)

Nos deux enseignants «à progrès élevé» semblaient avoir adopté des styles d'enseignement pour les mathématiques qui différaient passablement de ceux de leurs collègues qui enseignaient des matières équivalentes à des élèves présentant des niveaux de rendement initial équivalents. Par exemple, les enseignants «à progrès scolaire faible» passaient beaucoup de temps à des leçons ou à des parties de leçons réservées à la révision; ils passaient moins de temps à surveiller le travail individuel et plus de temps à la direction de la classe; de plus, ils passaient de plus longues périodes sans se mêler aux élèves. Par contre, lorsqu'ils préparaient leurs leçons et qu'ils les donnaient, les enseignants «à progrès élevé» insistaient de façon spéciale sur la nouvelle matière plutôt que de revoir de la matière déjà vue. Même si leurs propres pratiques d'enseignement étaient différentes, ces différences s'expliquent peut-être justement par le talent de ces deux enseignants à diagnostiquer les habiletés particulières de leurs élèves et à adopter des stratégies d'enseignement convenant bien à ces élèves-là.

Bref, les analyses secondaires des données provenant de l'une des études de l'Ontario soutenaient l'hypothèse que les intentions des enseignants importent davantage que leur fréquence moyenne d'utilisation de types d'interactions particuliers, pour qui cherche à comprendre les progrès scolaires des élèves. Les intentions des enseignants se reflétaient dans leur façon de couvrir la matière et l'orientation vers des buts éducatifs précis qu'ils réussissaient à imprimer à leurs classes. Comme ils étaient différents dans le genre de leçons qu'ils préparaient (leur façon d'utiliser le travail individuel, par exemple), nous n'avons pu dégager aucune «façon idéale d'enseigner». Les deux enseignants

étaient efficaces, à notre avis, car ils savaient préparer des cours bien adaptés à leurs élèves et réalisaient leurs intentions.

La deuxième analyse secondaire des données de l'ÉCSC résulta de l'analyse des données internationales, plus particulièrement, le regroupement des comportements d'interactions pour toutes les occasions, toutes les leçons et tous les enseignants. Les résultats de nos analyses initiales nous ont permis de nous demander si les regroupements éliminaient les intentions des enseignants au lieu de supprimer les erreurs des observateurs. Nous nous sommes dits que des tests sur les liaisons entre les comportements des enseignants au niveau des leçons et les comportements (que ce soit apprendre ou participer) des élèves à ce même niveau nous livreraient des conclusions plus significatives et se prêtant davantage à l'interprétation.

Pour ces raisons, Mandeville (1984) utilisa les leçons (plutôt que les classes) comme unités d'analyse dans une analyse secondaire des données internationales. Il testa d'abord les corrélations entre les comportements observés chez les enseignants (INST et ICM) et la participation (prise comme une «procuration» au niveau de la leçon, du rendement des élèves) observée chez les élèves. Il trouva premièrement qu'il y avait d'exceptionnelles variations entre les enseignants en ce qui concernait à la fois la direction et l'ampleur de ces liaisons lorsque la leçon constitue l'unité de regroupement. Par exemple, la corrélation entre le pourcentage de temps qu'un enseignant ou une enseignante passait à faire un cours magistral et le pourcentage d'élèves qui participaient pendant cette activité était négative (-0,65); pour toutes les leçons enseignées par cet enseignant ou cette enseignante, le taux de participation des élèves le plus élevé fut atteint pendant une leçon au cours de laquelle on observa l'utilisation la plus réduite de cours magistral. Cependant, pour un autre enseignant ou une autre enseignante dans le même pays, la corrélation entre le temps passé en cours magistral et le pourcentage d'élèves qui participaient était positive (+0,87), et la leçon pendant laquelle on avait observé que l'enseignant ou l'enseignante consacrait le moins de temps à l'enseignement magistral était la leçon dans laquelle le taux de participation des élèves était le plus bas. Bref, ce qui semblait «marcher» pour un enseignant ou une enseignante cherchant à faire participer les élèves ne «marchait» pas nécessairement pour un(e) autre.

Deuxièmement, Mandeville (1984) trouva que si on catégorisait les comportements «typiques» des enseignants en utilisant des corrélations moyennes (pour tous les enseignants faisant partie d'un échantillon), les résultats produiraient dans plusieurs cas des modèles intéressants. C'est à dire qu'avec les enseignants

«typiques», les taux de participation des élèves étaient supérieurs dans les leçons ayant les caractéristiques suivantes : l'enseignant ou l'enseignante fait appel à l'enseignement magistral (particulièrement avec du matériel de soutien) en grand groupe; encourage la participation de tout le grand groupe; utilise diverses questions et les redirige lorsque nécessaire; réagit aux réponses des élèves aux questions qu'il pose, les encourage; indique si ces réponses sont correctes ou non; et se livre à un minimum d'activités de direction de classe pendant les leçons. Ces caractéristiques au sens large donnaient des corrélations relativement conséquentes pour tous les pays avec les taux de participation scolaire des élèves, bien que de grandes variations aient existé dans les corrélations parallèles pour les enseignants individuels. Cependant, Mandeville trouva tout de même une catégorie de comportements des enseignants pour laquelle il y avait des liaisons logiques avec la participation des élèves pour toutes les leçons et tous les enseignants. Il s'agit de la direction de classe. Les données montrèrent clairement que les élèves étaient moins attentifs à la matière scolaire pendant les leçons où les activités de discipline et de direction de classe sont fréquentes, quels que soient l'enseignant ou l'enseignante et le pays.

Même si les résultats de ces analyses «moyennes» concordaient jusqu'à un certain point avec les conclusions des autres chercheurs «processus-résultats», Mandeville (1984) conseilla, pour deux raisons, de se garder des généralisations à partir de ces conclusions. Premièrement, les conclusions ne s'appliquent pas à tous les enseignants (sauf peut-être l'importance de limiter le plus possible les activités de direction de classe). Deuxièmement, ces conclusions ne peuvent être perçues comme normatives, puisqu'on ne peut leur attacher aucune causalité. Mandeville conclut : «Il se pourrait bien que d'autres variables influentes, comme les caractéristiques des leçons, soient les causes d'un grand nombre de liaisons qui ont été identifiées» (1984, p. 161). (traduction libre)

Pour donner suite aux hypothèses que nous commençons à formuler sur l'importance pour les enseignants de préciser leurs intentions et bien planifier leur exécution, Mandeville (1984) conçut aussi une façon d'examiner la nature des leçons. À partir de données de trois pays, il entreprit une analyse descriptive de la séquence d'activités de chaque leçon (e.g. pendant une leçon de trois activités, l'enseignant ou l'enseignante pourra donner de l'enseignement magistral, donner du travail individuel, puis revenir à l'enseignement magistral). Il trouva un grand nombre de séquences différentes d'activités pour tous les enseignants. Dans un pays, par exemple, les 323 leçons observées se caractérisaient par 158 séquences uniques d'activités.

Dans l'échantillon anglophone de l'Ontario, Mandeville trouva que 17 % des 173 leçons ne comportaient qu'une activité. De façon surprenante, pour un tiers de ces 30 leçons à une seule activité, les élèves devaient faire du travail individuel pendant toute la période. Le reste des 20 leçons à une seule activité comportaient respectivement de l'enseignement magistral, ou de la révision, ou des discussions. La structure de leçon la plus répandue (28 %) contenait deux parties. Ces leçons étaient de deux types : (1) une partie discussion suivie d'un travail écrit individuel ou (2) une partie enseignement magistral suivie de travail écrit individuel. Parmi les 93 leçons restantes, il y avait 54 structures uniques. La plupart contenaient entre deux et quatre parties; la leçon la plus complexe comportait quatre parties de révision, trois parties d'enseignement magistral et s'achevait par du travail écrit individuel. Pour terminer, il est intéressant de noter que des 173 leçons, 65 % commençaient par une première partie (enseignement magistral ou discussion) qui durait en moyenne de 50 % à 60 % du cours. Environ 75 % des leçons se terminaient par du travail individuel et ces devoirs occupaient en moyenne un tiers de la période.

Les analyses de Mandeville donnent du poids à la généralisation que les comportements des enseignants, si on les extrait de leur contexte et qu'on les regroupe, sont non seulement instables, mais se relient également de façon différente à la participation des élèves selon les différents enseignants. Les deux réanalyses de l'Ontario et celles de Mandeville appuient l'idée que les éléments déterminants importants de l'efficacité des enseignants sont le caractère opportun et le caractère approprié des comportements des enseignants plutôt que leur fréquence. Enfin, les deux analyses confirment les conclusions sur l'importance de réduire le temps consacré en classe à la direction et de favoriser la création d'une orientation vers des buts éducatifs.

Conclusions

La description de l'étude sur le contexte de la salle de classe de l'AIERS a montré le grand soin et la rigueur qui ont caractérisé notre tentative d'utiliser un modèle de recherche «processus-résultats» afin de dégager des connaissances sur les comportements des enseignants efficaces. Malgré nos efforts, l'étude n'a pas réussi à produire ces connaissances. Les résultats des analyses internationales, lorsqu'on les a considérés dans leur ensemble, ne nous ont pas incités à aller de l'avant avec le projet initial, qui était de se servir des conclusions de l'étude corrélationnelle pour mettre au point des programmes de formation des enseignants et s'engager dans une phase expérimentale de l'étude de l'AIERS. Par conséquent, l'étude sur le contexte de la salle de classe prit

fin en même temps que sa phase corrélationnelle.

Bien que l'ÉCSC n'ait pas atteint tous ses buts, elle a tout de même fourni une base pour mieux comprendre ce qui se produit dans les salles de classe du monde entier. Anderson (sous presse) fait remarquer :

> Connaître les différences sur la matière vue, la variabilité dans le temps des comportements des enseignants et le fait que les enseignants ont recours à un nombre relativement restreint d'activités et de comportements pendant l'année scolaire, est fort utile pour ceux qui essaient de concevoir des études novatrices et plus productives sur le contexte des salles de classe et les pratiques éducatives. Les connaissances sur le rôle central des élèves (à la fois leurs antécédents et leurs perceptions de ce qui se produit effectivement en classe) dans le processus d'apprentissage donnent à penser qu'il faut inclure les élèves dans toutes les conceptualisations de l'enseignement en classe. (traduction libre)

Nos efforts visant à mener des analyses secondaires des données furent jusqu'à un certain point plus satisfaisants en termes des indications qu'ils nous ont données sur l'efficacité de l'enseignement. Cependant, nous étions limités par les «outils» de recherche utilisés dans l'étude internationale. Par exemple, nous avons commencé à croire que les intentions des enseignants étaient primordiales pour comprendre l'efficacité de l'enseignement; pourtant, nous disposions de peu de données qui nous auraient permis de tester cette hypothèse sur ces intentions. Nous avons commencé à nous rendre compte que les activités et les événements se déroulant en classe se transformaient en profondeur en fonction des perceptions et des sentiments des élèves; encore une fois, nous n'avions pas suffisamment de données pour vérifier cette hypothèse. Chacune de ces hypothèses, de même que les raisonnements qui nous y ont menés, indiquaient la nécessité de concevoir autrement le processus enseignement-apprentissage en classe. Le modèle «processus-résultats» utilisé pour la recherche sur l'efficacité des enseignants comporte trop de lacunes pour que nous puissions franchir la distance entre les conclusions des recherches et des connaissances utiles pour améliorer l'enseignement et l'apprentissage.

III. UN NOUVEAU MODÈLE D'EFFICACITÉ DES ENSEIGNANTS FONDÉ SUR LA CONCEPTION QUE L'ENSEIGNEMENT EST UN ENSEMBLE DE TÂCHES

Mes collègues et moi avons longuement réfléchi sur les expériences et les frustrations que nous a fait vivre pendant quatre ans l'étude sur le contexte de la salle de classe de l'AIERS. Après coup, les critiques sur les études processus-résultats qui ont été exprimées jusqu'en 1978 nous sont revenues à l'esprit. Une généralisation s'est dégagée de nos réflexions individuelles et collectives : il faut remettre en question la validité des suppositions qui sous-tendent le modèle (ou paradigme) des processus et des résultats.

Il est instructif de passer en revue les suppositions sous-tendant le modèle processus-résultats (ainsi que leurs critiques exprimées dans la documentation) et que nous avons transposées dans nos expériences dans le cadre de l'étude de l'AIERS. En association avec Anderson (1984), nous avons identifié trois suppositions fondamentales qui ont influencé à la fois le contenu et la forme de l'étude de l'AIERS.

Supposition 1. Il est possible de dériver empiriquement des recommandations pour l'amélioration de l'éducation.

Cette supposition signifie qu'un cadre de travail théorique ou conceptuel n'est pas a priori nécessaire comme point d'appui pour concevoir des outils à même d'améliorer les conditions éducatives. Les chercheurs peuvent se dispenser d'un tel cadre de travail pour tout simplement se pencher sur ce qui «marche» sous des conditions normales en classe et utiliser les résultats de ces recherches empiriques pour concevoir du matériel de formation pour les enseignants.

Supposition 2. Un modèle d'étude linéaire à deux phases (corrélationnelle et expérimentale) convient à la recherche destinée à améliorer les conditions éducatives.

Cette supposition découle de la première. La phase corrélationnelle est nécessaire pour identifier ce qui «marche» pour les «bons» enseignants; le «traitement» expérimental (un programme de formation des enseignants) est établi à partir de l'étude naturelle; et la phase expérimentale s'impose pour vérifier si le «traitement» fonctionne pour les enseignants et si les comportements «enseignés» se traduisent par un meilleur apprentissage par les élèves.

On suppose que ce modèle en deux phases est une façon utile de considérer la recherche axée sur l'amélioration des conditions éducatives car, par définition, les enseignants apprendront à utiliser les comportements que les «bons» enseignants utilisent déjà. On ne demandera pas aux enseignants d'utiliser des comportements «ésotériques» ou des pratiques qui ne sont «valables qu'en théorie». Autrement dit, comme l'ont écrit Gage et Fordham (1980, p. 4) dans leur proposition pour l'ÉCSC de l'AIERS, «Il est plus facile d'apporter à l'école des changements mineurs que majeurs.» (traduction libre)

Supposition 3. Le paradigme processus-résultats convient pour étudier l'enseignement et l'apprentissage en classe.

Encore une fois, cette supposition découle des deux autres. Exprimé simplement, le paradigme processus-résultats suppose que les comportements des enseignants et leurs pratiques (processus) influencent le rendement des élèves (résultats). Intuitivement, il est logique que la recherche axée sur l'amélioration de l'éducation s'arrête sur ce que les enseignants font en classe et sur les liens entre ce qu'ils font et ce que les élèves apprennent. Cette approche est beaucoup plus utile que d'examiner les traits des enseignants (caractéristiques professionnelles et personnelles et compétence des enseignants). De plus, ce paradigme est une solution intéressante pour remplacer le paradigme ressources-résultats car nous croyons tous que les enseignants font une différence.

Remise en question de la validité des suppositions processus-résultats

Selon nos expériences dans l'étude de l'AIERS, les trois suppositions ci-dessus, tout intéressantes qu'elles soient, ne semblent pas être des sources d'inspiration indiquées et adéquates pour la recherche axée sur l'amélioration de l'éducation. D'autres chercheurs partagent notre opinion à ce sujet. Par exemple, Doyle (1978) a émis de sérieuses réserves sur le paradigme processus-résultats précisément à cause de ces suppositions fondamentales. Il a écrit que le paradigme renferme peu de principes explicatifs clairs qui permettent d'inspirer le choix des variables ou l'interprétation des résultats. Sa structure à deux facteurs ne tient pas compte des comportements d'apprentissage des élèves et des autres variables qui pourraient contribuer à expliquer comment les comportements des enseignants influencent les résultats obtenus par les élèves dans leur apprentissage. En pratique, a écrit Doyle, le choix des variables semble s'effectuer sur la base de préférences personnelles ou de critères strictement empiriques telle l'ampleur des coefficients de corrélation. Bien que les commen-

taires sommaires de Rosenshine (1976) et d'autres soient utiles, Doyle faisait valoir que de tels résumés n'expliquaient pas pourquoi il faudrait s'attendre à ce que de tels comportements des enseignants ou variables présentent des liens avec le rendement. «En l'absence de propositions explicatives formelles, il est difficile d'interpréter des conclusions contradictoires ou de choisir des voies qui pourraient se révéler intéressantes à explorer (p. 167).»

Nous avons fait face à ces problèmes directement dans l'étude de l'AIERS. Notre longue liste de variables ressemblait à une énumération des comportements préférés de différentes personnes. Les comportements que nous avons étudiés étaient si nombreux et représentaient de si «petites tranches» de la vie en classe qu'ils ne permettaient en aucune manière de reconstituer le contexte global de la classe. En l'absence d'un cadre de travail théorique ou conceptuel pour choisir les variables à inclure dans notre énumération, nous avons eu du mal à comprendre pourquoi nous n'arrivions à presque rien avec les catégories éducatives qui prédisaient le rendement des élèves au post-test. Les analyses secondaires mentionnées à la fin de la partie II représentent nos efforts d'essayer d'autres façons de reconstruire le contexte de la salle de classe et d'examiner les processus qui pourraient influencer l'apprentissage des élèves. Cependant, comme nous l'avons noté, nous étions limités par la nature des données issues de notre modèle «athéorique».

Une autre critique du paradigme processus-résultats est qu'il met en lumière presque essentiellement les variables des enseignants et, par surcroît, seulement des aspects choisis du travail des enseignants. Doyle (1978) a critiqué deux de ces aspects en particulier. Le premier de ces aspects, la fréquence d'un comportement, était causée par l'utilisation d'un horaire d'observation à inférence faible. De tels systèmes enregistrent des exemples discrets de comportements particuliers. Ceci implique que la récurrence d'un comportement détermine ses effets et que plus elle est forte, mieux c'est. Cependant, la supposition que la fréquence détermine les effets semble n'avoir que peu de fondement empirique ou conceptuel. Doyle a cité les études sur les travaux de Brophy et Evertson (1974) et Soar (1972) qui auraient suggéré que dans le cas de certains comportements, c'est le niveau optimal, plutôt que la quantité absolue, qui détermine les effets sur l'apprentissage des élèves. Doyle a également noté que pour certains comportements, le moment opportun importe davantage que la fréquence d'emploi. Mandeville (1984), dans ses analyses des données de l'ÉCSC de l'AIERS confirme certainement aussi cette critique.

Les études processus-résultats ont aussi été critiquées à cause de leur accent

sur les comportements des enseignants qui sont des comportements stables et peuvent s'appliquer de façon générale dans tous les contextes, pour tous les genres d'apprenants et de tâches. Lorsqu'une variable cause différents effets à des moments différents ou avec des groupes d'élèves différents, elle n'offre pas de grande valeur pratique pour la conception des programmes de formation des enseignants. Alors, à l'extrême, le chercheur processus-résultats finit par rechercher des «lois» sur l'efficacité des enseignants qui, selon Doyle (1978, p. 169), sont «hors-contexte, à l'épreuve des enseignants et même des élèves».

Comme le démontrent les résultats de l'étude de l'AIERS, les comportements des enseignants tendent à être instables (i.e. ils changent beaucoup d'une leçon à l'autre, selon les occasions, pour les enseignants individuels). Cependant, le paradigme processus-résultats présuppose une grande stabilité dans le comportement de l'enseignant ou enseignante; sinon, il ne serait pas possible de relier les observations effectuées sur une courte période de temps aux résultats mesurés à la fin de l'étude. Non seulement cette supposition fut-elle contredite empiriquement dans l'étude de l'AIERS, mais elle peut également être remise en question en vertu des intentions des enseignants qui les incitent à adapter leurs comportements aux événements survenant en classe. Par exemple, Medley (1982, pp. 1898-99), conclut :

> Si par exemple, un enseignant ou une enseignante posait à dessein plus de questions lorsqu'il ou elle présente une nouvelle unité que lorsqu'il ou elle en révise une, et qu'il ou elle encourageait les élèves à poser plus de questions pendant la révision que pendant la présentation de matière nouvelle, ce fait ne serait pas détecté car une moyenne de ces variations serait établie, et ce, pour les deux types de situations. Il est évident que les changements intentionnels de styles d'enseignement peuvent constituer des éléments importants dans la capacité de rendement d'un enseignant ou d'une enseignante; le talent d'un enseignant ou d'une enseignante à changer de styles d'enseignement pourrait bien être un facteur déterminant dans l'efficacité de cet enseignant ou cette enseignante. Les études processus-résultats ont presque complètement négligé cet élément important du rendement des enseignants. (traduction libre)

Enfin, les résultats de l'étude de l'AIERS suggèrent que les enseignants disposent d'un répertoire limité de comportements qu'ils utilisent dans le processus des interactions en classe. Par exemple, nous n'avons presque jamais observé d'enseignants en train de donner des exemples ou de vérifier si les élèves comprenaient une explication. Dans une étude plus récente, Anderson (1986b), a constaté que 88 enseignants au secondaire présentaient rarement des objectifs éducatifs clairs au début d'une période de classe; ils faisaient rarement comprendre aux élèves qu'ils étaient capables d'apprendre et qu'on attendait d'eux qu'ils apprennent la matière présentée; ils encourageaient rarement les élèves à

faire des efforts pour apprendre; et ils vérifiaient rarement si les élèves avaient fait ou révisé les devoirs. Le paradigme processus-résultats exigerait que ces comportements peu fréquents soient exclus des programmes de formation des enseignants dérivés de la phase naturelle corrélationnelle. Good, Grouws, et Ebmeier (1983) conclurent que les enseignants de mathématiques, sur lesquels avaient porté leurs études, consacraient rarement du temps de classe à une partie «développement» d'une leçon, même si les éducateurs en mathématiques insistent sur l'importance de prendre le temps de développer les concepts ou les idées que les élèves doivent assimiler. Encore une fois, si les éducateurs s'étaient contentés de laisser leur étude «naturelle» dicter ce qu'ils devaient faire comme efforts de formation expérimentale, les enseignants n'auraient pas été encouragés à planifier des leçons comprenant une partie «développement».

À partir de nos réflexions sur nos propres conclusions et celles d'autres chercheurs, nous avons proposé, en collaboration avec Anderson (1984), trois autres suppositions susceptibles de guider la recherche orientée vers l'amélioration de l'efficacité des enseignants. Voici ces suppositions :

1. Il faut un cadre de travail théorique ou conceptuel pour comprendre et finalement améliorer le processus de l'enseignement et de l'apprentissage.

2. Le modèle des études expérimentales doit s'appuyer à la fois sur un cadre conceptuel et sur des connaissances empiriques acquises grâce à diverses études (i.e. des études expérimentales aussi bien que corrélationnelles).

3. Le paradigme processus-résultats est une simplification à outrance du monde de l'enseignement en classe et de l'apprentissage scolaire.

Nous croyons qu'il existe un besoin net de créer un nouveau paradigme ou modèle qui inspirera les recherches futures dans le domaine de l'enseignement et qui, par conséquent, permettra d'en arriver à des conclusions qui seront valides tout en étant utiles aux praticiens et aux responsables de la formation des enseignants. Ce modèle devrait tenir compte de la complexité de l'enseignement et devrait s'orienter à la fois vers la compréhension et l'amélioration de l'enseignement. Autrement dit, nous avons besoin d'un cadre qui nous donne un point de référence pour examiner et comprendre les intentions des enseignants, les raisons qui font que certains comportements ou pratiques d'enseignement conviennent bien à certains moments et ne conviennent pas à d'autres et, par conséquent, pourquoi les enseignants agissent différemment (de façon apparemment incohérente) d'une situation à l'autre. Il nous faut également un cadre conceptuel qui nous permette de tenir compte des variations propres à chaque enseignant au lieu de tenter de se débarrasser de cette variation en «clonant» les enseignants

par l'emploi de programmes de formation tout à fait hors contexte.

L'enseignement : un ensemble de tâches

Dans un numéro spécial de Evaluation in Education (Ryan et Anderson, 1984), nous avons, avec nos collègues, présenté un exemple de cadre conceptuel. Notre cadre ou modèle reposait sur la supposition que ce sont les accomplissements des enseignants, et non pas ce qu'ils sont ou font, qui les rendent efficaces à différents degrés. Dans les articles publiés dans ce numéro spécial, puis dans d'autres articles par la suite (voir Ryan, 1985 a et b; Anderson, 1985, 1986 a et b), nous avons adopté l'opinion que les bons enseignants sont capables de réaliser un ensemble complexe de fonctions ou de tâches. Ensemble, ces tâches produisent le comportement en classe voulu de la part des élèves, les incitent à participer activement à leur apprentissage et finalement, leur permettent d'atteindre des niveaux élevés d'apprentissage.

Cette idée commença à gagner des défenseurs à la fin des années soixante-dix et au début des années quatre-vingts, représentant des disciplines variées, dont la sociologie, la psychologie et la recherche et le développement en éducation (Anderson, 1986; voir aussi Rosenshine, 1983; Brophy et Evertson, 1978; Brophy, 1981; et Good, Grouws, et Ebmeier, 1983). Le travail combiné de ces chercheurs permit de mettre le doigt sur deux problèmes majeurs de la recherche basée sur la conception que l'efficacité des enseignants est reliée à leurs comportements. Le premier problème, c'est que les comportements des enseignants prennent tout leur sens dans le milieu même où ils se produisent; par conséquent, il faut diriger notre attention sur le caractère approprié de ces comportements, pas seulement à leur présence, leur absence, leur fréquence ou leur durée. Le deuxième problème est que la plupart des actions des enseignants en classe sont motivées par un but particulier et que les motifs des mêmes comportements peuvent changer selon les circonstances et les moments; par conséquent, il faut considérer les buts des enseignants et pas seulement la constance (ou le manque de constance) des comportements observables en classe.

La conception que l'enseignement est un ensemble de tâches nous permet de nous attaquer à ces deux problèmes majeurs. Une tâche se définit comme «une unité de travail discrète, organisée, comportant un début défini, une fin, et accomplie par quelqu'un pour réaliser les buts d'un travail» (Gael, 1983). Les tâches nous permettent de saisir pourquoi les enseignants font ce qu'ils font ou, autrement dit, à quoi servent leurs comportements (réf. Rosenshine, 1983; Brophy, 1981). Par exemple, les enseignants posent-ils des questions afin de

reprendre l'attention de certains élèves (une tâche de direction) ou afin de vérifier si les élèves comprennent un élément de matière nouvelle (une tâche d'enseignement)? Les tâches nous donnent aussi l'occasion de nous pencher sur le caractère approprié des comportements des enseignants en mettant l'accent sur la séquence des tâches des enseignants pendant les trimestres, semaines, jours d'enseignement. Dans notre modèle d'enseignement, nous supposons que certaines tâches doivent être réalisées à certains moments au cours de l'année scolaire. Par exemple, si les tâches de fixer et de clarifier les règles et les habitudes régissant le comportement des élèves ne s'établissent pas pendant les deux ou trois premières semaines de l'année scolaire, les enseignants devront consacrer plus de temps pendant l'année à établir et maintenir l'ordre, aux dépens des autres tâches éducatives.

Enfin, la conception que l'enseignement est un ensemble de tâches nous donne un cadre utile (et valide, croyons-nous) pour tenter d'améliorer l'enseignement. Pour les superviseurs des enseignants, par exemple, il est beaucoup plus facile de juger si les tâches d'un enseignant ou d'une enseignante ne sont pas appropriées que de dire que ce sont ses comportements qui laissent à désirer. Pour les enseignants, le fait de terminer une tâche définie porte en soi sa récompense et, par conséquent, rend cette tâche plus agréable. Nous reviendrons plus tard aux implications de ce modèle dans la pratique. D'abord, discutons de la conceptualisation des tâches des enseignants que nous avons développée en collaboration avec nos collègues. Cette discussion provient de Ryan (1985).

Cadre des principales tâches des enseignants

Le point de départ de notre conceptualisation de l'enseignement reproduit dans l'illustration 1 fut notre supposition que la pratique professionnelle de l'enseignement exige la réalisation d'une grande quantité de tâches diverses, dont plusieurs doivent être exécutées en dehors de la classe. Voyez le résumé qui suit :

> Dans la plupart des pays, un enseignant ou une enseignante, parfois un(e) assistant(e), ainsi que de 20 à 40 élèves, occupent une salle de classe pendant 5 à 6 heures par jour, pendant 180 jours. Sur une période approximative de 40 semaines, le travail principal d'un enseignant ou d'une enseignante est de développer chez les élèves une gamme de buts cognitifs et affectifs. Afin de réaliser ces buts à long terme, les enseignants doivent planifier et organiser le programme de l'année scolaire, choisir et mettre sur pied des activités qui soutiendront la participation des élèves, gérer les interactions sociales et éducatives de façon ordonnée et sans heurts, et prendre en considération les différences entre les élèves de manière à ne pas ennuyer les élèves apprenant plus vite et ne pas frustrer ceux qui apprennent

> moins vite. Les enseignants doivent accomplir leur travail en harmonie avec les attentes de la collectivité, en étant soumis à des contraintes administratives et organisationnelles et composer avec des interruptions assez prolongées du processus d'apprentissage scolaire pendant l'année. Parfois, ils sont aux prises avec des élèves fort récalcitrants, qui refusent de participer aux activités en classe (Burns, 1984, p. 96). (traduction libre)

La documentation sur la planification et la prise de décisions de l'enseignant ou enseignante nous a aidés à organiser les tâches variées des enseignants. Nous avons trouvé que les responsabilités professionnelles des enseignants se caractérisaient idéalement en termes de décisions dont la nature se modifie selon deux facteurs : le moment où les décisions sont prises et les demandes contextuelles qui commandent des décisions sur le contenu et les processus recherchés (que nous avons nommés unités structurales). Nous nous sommes dit que les décisions prises à différents moments et au sujet de choses différentes se prenaient également pour des durées précises du temps d'enseignement en classe. De plus, ces durées voulues semblaient correspondre à l'organisation et à la structure de l'année scolaire. Enfin, nous avons pensé que le travail des enseignants, sur le plan des points de décision et des unités structurales, définirait différents types de tâches des enseignants. Les tâches sont des buts à atteindre par les enseignants et des séries de comportements menant à la réalisation de ces buts; en conséquence, nous avons perçu les tâches comme faisant le lien entre les comportements des enseignants et leurs intentions ou propos.

Opportunité des décisions

La première dimension de l'illustration 1, l'opportunité des décisions, est fondée sur la distinction de Jackson (1962) entre les décisions préactives et interactives de l'enseignant ou enseignante. L'enseignant ou l'enseignante prend des décisions préactives à l'extérieur de la salle de classe, souvent seul(e); ce sont des décisions de planification des activités qui auront lieu en classe pendant une période de temps à venir. Il existe deux sortes de décisions préactives, à long terme et à court terme; elles varient selon la planification particulière et la durée prévue d'utilisation de cette planification dans le contexte de la classe. L'enseignant ou l'enseignante prend les décisions interactives dans la salle de classe. Ces décisions jouent un rôle dans le rythme des événements éducatifs et des interactions avec les élèves et découlent des jugements de l'enseignant ou enseignante sur les comportements et le déroulement des activités à mesure que progresse la leçon planifiée.

ILLUSTRATION 1. Conceptualisation schématisée de l'enseignement en termes de décisions, d'unités structurales et des tâches des enseignants.

TYPES DE TÂCHES DES ENSEIGNANTS

OPPORTUNITÉ DES DÉCISIONS	UNITÉ STRUCTURALE	TÂCHES ÉCOLOGIQUES				
Décisions préactives à long terme	Cours ou période scolaire	Choix des buts, sujets, matériel, activités	Planif. périodes de temps allouées	Planification constitution des groupes	Mise au point de règles et habitudes de présentation de la matière	Mise au point de règles et habitudes de direction de la classe

OPPORTUNITÉ DES DÉCISIONS	UNITÉ STRUCTURALE	TÂCHES DE SOUTIEN ET DE DÉVELOPPEMENT ÉDUCATIFS		
Décisions préactives à court terme	Unité d'apprentissage	Choix du contenu et des objectifs	Développement des unités d'apprentissage et des leçons	Développement des tests et des devoirs des unités
	Leçon	Choix des objectifs de la leçon	Développement des segments de la leçon	Développement des devoirs de la leçon
	Segment de la leçon	Décision sur le but du segment de la leçon	Décision sur le format de l'activité	Décision sur les devoirs du segment

OPPORTUNITÉ DES DÉCISIONS	UNITÉ STRUCTURALE	TÂCHES D'ENSEIGNEMENT					TÂCHES DE DIRECTION		
Décisions inter-actives	Épisode	Orienter les élèves	Enseigner	Surveiller, évaluer les progrès	Donner réactions	Corriger erreurs d'app.	Clarifier règles	Maintenir rythme des activités	Punir les comportem. indésirables

Unités structurales

Les unités structurales, qui forment la deuxième dimension de l'illustration 1, sont les «choses» sur lesquelles les enseignants prennent des décisions et, par conséquent, elles sont apparentées de près à la dimension temporelle. En ordre décroissant par limite de temps, il y a cinq unités structurales : le cours, l'unité d'apprentissage, la leçon, le segment de la leçon et l'épisode. Chaque unité inférieure s'imbrique dans l'unité supérieure suivante; par exemple, les leçons s'imbriquent dans les unités d'apprentissage.

L'unité la plus grande, le cours, est la limite temporelle naturelle selon laquelle on organise le programme scolaire, structure les activités éducatives et évalue formellement l'apprentissage des élèves. Le cours est constitué d'unités d'apprentissage, une construction à la base de l'organisation du programme. Lorsqu'ils prennent leurs décisions sur le contenu, les activités et l'évaluation des unités d'apprentissage, les enseignants planifient généralement leur enseignement pour une ou deux semaines. Dans les unités d'apprentissage, s'imbriquent les leçons, les unités de planification et de structure de base de la classe. En planifiant leurs leçons, les enseignants préparent les objectifs, les séquences d'activités et l'évaluation informelle qui prendront place de façon typique dans une période ou un jour de classe. Les segments de leçons, à leur tour, procurent le format ou la structure des activités et donnent forme aux rôles prévus des enseignants et des élèves pour un certain nombre de minutes du temps de la leçon. Par exemple, dans les classes de mathématiques, les enseignants préparent souvent une séquence de trois segments pour une leçon : révision, développement d'un concept ou d'une opération et devoir individuel.

Les épisodes s'imbriquent dans les segments de leçons; ils consistent en séquences limitées d'interactions ou de comportements. On peut retrouver un grand nombre de ces épisodes à l'intérieur de chaque segment d'une leçon. Le concept d'épisodes de comportements provient du travail de psychologues écologistes (Barker, 1963; Gump, 1964) et nous le voyons comme représentant la plus petite unité significative de comportements observables en classe. Les épisodes se caractérisent à la fois par un comportement ou une action et par le contexte situationnel (c'est à dire la structure de l'activité) dans lequel le comportement se produit.

Il est important de noter que chaque unité structurale inférieure se définit par l'unité supérieure suivante dans laquelle elle s'imbrique. Ainsi, le sens d'un épisode de comportements dépend du contexte (segment) de l'activité à l'inté-

rieur duquel elle se produit. Un segment tire son sens de la nature de la leçon dans laquelle il a lieu. Le sens d'une leçon doit dériver de l'unité d'apprentissage dont elle ne représente qu'un élément, dans une séquence planifiée d'enseignement.

Types de tâches des enseignants

La troisième dimension montrée dans l'illustration 1, les types de tâches des enseignants, donnent des moyens d'identifier les unités de travail (tâches) reliées que les enseignants doivent exécuter pour assumer leurs responsabilités fonctionnelles. Nous avons identifié quatre principaux types de tâches des enseignants; chacun s'insère dans les décisions à prendre à différents moments au cours de l'année scolaire et concerne les différents éléments (unités structurales) du contenu et des processus éducatifs voulus.

Nous avons nommé le premier ensemble tâches écologiques car elles servent à établir les contraintes générales relatives au programme, temporelles et sociales, dans le cadre desquelles se produit l'enseignement. À partir de la revue de Burns (1984) des études ethnographiques sur les enseignants, nous avons identifié cinq sous-tâches devant être exécutées au début de l'année scolaire : choix des buts, des sujets, du matériel et des activités du cours; préparation des périodes de temps allouées aux sujets; décisions sur la pertinence, le moment et la manière de constituer des groupes d'élèves pour l'étude de sujets de matière; développement de règles et d'habitudes de présentation de la matière; et développement de règles et d'habitudes de direction de classe. Chacune de ces sous-tâches exige des décisions préactives à long terme en rapport avec la simple unité structurale du cours.

Les décisions préactives à long terme que les enseignants prennent en accomplissant des tâches écologiques fournissent des lignes directrices générales pour la planification subséquente des éléments constitutifs du cours. Il importe donc de reconnaître que les mandats du système scolaire ou de l'instance provinciale peuvent retirer ou imposer la responsabilité des enseignants pour les deux premières sous-tâches (choix du contenu du programme, du matériel didactique et des périodes de temps allouées). De même, les structures et les politiques organisationnelles des districts et des écoles (par exemple, nombre d'élèves par classe, règles sur le placement des élèves, horaire des cours) peuvent limiter le choix des enseignants en ce qui a trait à la constitution des groupes et les obliger à utiliser des règles et des procédures d'enseignement et de direction particulières.

En suivant ces lignes directrices générales, les enseignants doivent accomplir une quantité de sous-tâches particulières qui, réunies, fournissent le fondement de l'enseignement réel en classe, conditionné par les exigences du programme et de l'enseignement. Nous les avons appelées tâches de soutien et de développement éducatifs. Bien que l'accomplissement de chacune de ces sous-tâches exige des décisions préactives à court terme, trois différentes unités structurales entrent en jeu : les unités d'apprentissage (les éléments constitutifs d'un cours), les leçons (les éléments constitutifs des unités d'apprentissage) et les segments de leçons (les éléments constitutifs des leçons).

Trois sous-tâches de soutien et de développement éducatifs doivent être accomplies pour chacune des trois unités structurales. Ainsi, chaque rangée de ce domaine de tâches dans l'illustration 1 attire l'attention sur une catégorie structurelle distincte de planification par l'enseignant ou l'enseignante. Chaque colonne du domaine de tâches met en relief un des trois éléments de tâches de planification de l'enseignant ou enseignante : contenu ou objectifs des programmes, activités et formats éducatifs, et devoirs et tests des élèves.

Notre esquisse des tâches de soutien et de développement éducatifs reflète le travail des chercheurs et des concepteurs en éducation adaptative et est en accord avec l'approche de la maîtrise des unités d'apprentissage appliquée à la planification des programmes et de l'éducation (Anderson et Jones, 1981; Anderson et Block, 1976; Ryan et Schmidt, 1979). Les systèmes de maîtrise des unités d'apprentissage exigent que les enseignants s'occupent systématiquement de chacune des tâches préactives à court terme figurant dans l'illustration 1. Pour chaque unité d'apprentissage, par exemple, il faut développer une séquence de leçons afin d'orienter les élèves vers les objectifs voulus, d'enseigner en fonction du contenu, de faire participer les élèves à des activités appliquant le contenu, d'évaluer leur apprentissage, de leur donner des réactions sur cet apprentissage et de corriger leurs erreurs d'apprentissage. On pourra développer une leçon entière à une seule fin (par exemple, faire participer les élèves à des activités pratiques), mais le facteur important est que le but de la leçon ne se comprenne que dans le contexte de plans pour l'unité prise en tant qu'entité. De plus, les enseignants peuvent modifier la programmation de leurs leçons en fonction des événements se produisant en classe et des réactions des élèves (par exemple, en accordant une période supplémentaire pour l'apprentissage ou des exercices supervisés), mais ils conservent la responsabilité de mener à bien toutes les sous-tâches prévues dans la séquence de leçons planifiées pour l'unité.

Les responsabilités fonctionnelles des enseignants dans l'exécution de leurs décisions préactives à court terme se définissent en deux groupes reliés de sous-tâches, les tâches d'enseignement et les tâches de direction. Les tâches d'enseignement servent à assurer la participation des élèves aux activités planifiées pour la leçon et aux tâches d'apprentissage, et les tâches de direction servent à la fois à maintenir le rythme d'enseignement et la participation et la coopération des élèves aux activités planifiées de la leçon. Les deux genres de tâches obligent les enseignants à prendre des décisions interactives en réponse aux actions et réactions des élèves, minute par minute. Ces deux sortes de sous-tâches s'accomplissent en référence aux unités structurales des épisodes de comportements.

Encore une fois, la revue de documentation de Burns (1984) servit de base pour identifier les sous-tâches nécessaires à la réalisation des tâches de direction et d'enseignement en classe. En analysant un certain nombre d'études à grande échelle de type «processus-résultats» ou expérimentales sur le terrain et de revues de ces études, Burns a découvert qu'il y avait un fort consensus pour un modèle général d'enseignement efficace. On appelle souvent ce modèle ou ensemble de comportements des enseignants l'enseignement direct, et il consiste en cinq sous-tâches d'enseignement reliées : orienter les élèves, enseigner aux élèves, surveiller et évaluer leurs progrès, leur donner des réactions, et corriger leur mauvais apprentissage quand cela se produit. Il y a un parallèle très proche entre cet ensemble de tâches de l'enseignant ou enseignante efficace en classe, identifié par la recherche sur l'enseignement, et les tâches préactives efficaces pour planifier une séquence de leçons à l'intérieur d'une unité d'apprentissage, qui ont été identifiées par le travail des concepteurs en éducation adaptative.

La documentation sur la direction de la classe nous a aidés à identifier trois sous-tâches dont la réalisation contribue à la création d'une atmosphère propice à un enseignement efficace, à l'intérieur de la salle de classe. Voici ces sous-tâches : clarifier les règles et les procédures, maintenir le rythme des activités des leçons, et sévir sans tarder contre les comportements indésirables.

Même s'il y avait un consensus parmi les auteurs qui ont procédé à des analyses des recherches en éducation au sujet des sous-tâches accomplies par les enseignants efficaces, Burns (1984) a trouvé qu'il y avait bien moins de consistance empirique concernant les comportements particuliers d'enseignement en classe qui étaient plus ou moins efficaces dans la réalisation d'une sous-tâche. Ce qui a semblé important, c'est l'apparition chez les enseignants d'un certain comporte-

ment propre à réaliser le but d'une sous-tâche, plus que l'apparition d'un comportement particulier. Par exemple, les enseignants devraient accomplir la sous-tâche de donner leurs réactions aux élèves. Cependant, il n'est pas bien clair si, dans l'accomplissement de cette tâche, il est plus efficace de faire des compliments ou des critiques. Nous soupçonnons que l'efficacité d'un comportement particulier rattaché à une tâche dépend du contexte de la salle de classe à ce moment.

Le concept d'équivalence fonctionnelle, emprunté à Smith et Geoffrey (1968), a aidé à comprendre que différents comportements d'enseignement peuvent parvenir aux mêmes fins, selon le contexte donné. De la même façon, un comportement peut répondre à différents buts, encore une fois, selon le contexte éducatif dans lequel il se produit.

En résumé, le travail d'un enseignant ou d'une enseignante s'accompagnait des responsabilités d'accomplir des tâches reliées à quatre principaux domaines fonctionnels : tâches écologiques, tâches de soutien et de développement éducatifs, tâches d'enseignement et tâches de direction. En réalisant les sous-tâches convenant aux quatre domaines, les enseignants doivent prendre des décisions variant en termes d'opportunité (décisions préactives à long et court termes, décisions interactives) et de référent structural (cours, unité d'apprentissage, leçon, segment de leçon, et épisode). Les décisions préactives à long terme (tâches écologiques) influencent les décisions préactives à court terme (tâches de soutien et de développement éducatifs) lesquelles à leur tour influencent l'accomplissement des tâches d'enseignement et de direction en classe. Lorsqu'ils suivent leurs plans, les enseignants doivent prendre des décisions interactives destinées à maintenir à la fois le rythme des activités éducatives planifiées et une forte participation des élèves. Ces expériences en classe permettent aux enseignants de voir les effets de leurs décisions de soutien et de développement éducatifs.

Le schéma de l'illustration 1 aide à montrer que les comportements des enseignants en classe ne peuvent se comprendre ou s'évaluer sans connaître les plans plus généraux de l'enseignement des unités et des leçons. De plus, si les enseignants n'ont pas accompli les tâches de planification préactives, on ne peut espérer qu'ils pourront réaliser les sous-tâches d'enseignement appropriées dans leur classe.

Implications de la conception que l'enseignement est un ensemble de tâches

Une des principales limites conceptuelles de l'opinion que l'enseignement est un ensemble de comportements (i.e. c'est ce que les enseignants font qui les rend efficaces), c'est que le seul point de référence demeure l'accomplissement de tâches d'apprentissage par les élèves. Les définitions utiles de Medley (1982), que vous retrouvez à la section 1 attirent l'attention sur ce point. Medley a noté que lorsqu'on juge de l'efficacité des enseignants selon les résultats obtenus par les élèves, un enseignant particulier ou une enseignante particulière peut fort bien être efficace avec un groupe d'élèves et ne pas l'être avec un autre groupe. Dans le jargon des chercheurs, les effets des enseignants sur différents groupes d'élèves sont «instables» et «incohérents» d'une année à l'autre. Rosenshine (1977) a même découvert que ces effets étaient instables d'un sujet d'étude à un autre, pour le même groupe d'élèves.

Évidemment, nous souhaitons tous découvrir des stratégies éducatives capables de se traduire, à long terme, par des améliorations soutenues de l'apprentissage des élèves. Cependant, nous croyons important de penser à la réalisation de diverses tâches par l'enseignant ou l'enseignante comme ayant davantage de résultats immédiats pour les enseignants. Par exemple, quand un enseignant ou une enseignante accomplit la tâche de décider des règles et des habitudes de comportements de sa classe (une tâche écologique), et quand il ou elle les applique et les clarifie à l'intention des élèves, (une tâche de direction de classe), il y a des «compensations» immédiates et continues : les élèves comprennent ces règles et ces habitudes et les adoptent; les activités en classe se succèdent harmonieusement; et il y a plus de temps pour l'enseignement. De la même façon, quand un enseignant ou une enseignante accomplit une tâche d'évaluation, le «résultat» ou la compensation est que l'enseignant ou l'enseignante acquiert une compréhension du niveau atteint de connaissances, d'habiletés ou de motivation chez les élèves. Cette compréhension peut servir à prendre différentes décisions importantes (e.g. Dois-je regrouper les élèves pour l'enseignement? Dois-je passer au prochain sujet? Dois-je donner une note d'échec à un élève?).

Si on accepte la perception que ce sont les accomplissements des enseignants qui les rendent efficaces, on peut voir l'enseignement comme une série de tâches, liées entre elles, dont chacune comporte des compensations particulières et identifiables pour l'enseignant ou l'enseignante. Comme le fait remarquer Anderson (1985), aucune des compensations des tâches d'un enseignant ou d'une enseignante individuel(le) n'est le rendement des élèves. La conception que l'enseignement est un ensemble de tâches suppose plutôt que «c'est seulement lorsque l'ensemble de tâches a été accompli sur une période de temps assez

considérable qu'il est possible de constater des changements appréciables dans le rendement des élèves» (Anderson, 1985, p. 21).

Grâce aux recherches entreprises adoptant diverses approches et perspectives au cours des dix dernières années, nous disposons de connaissances utiles sur les tâches des enseignants et les façons dont ils peuvent les accomplir efficacement. Si on les interprète de façon générale, les conclusions auxquelles sont parvenus les psychologues écologistes, les ethnographes, les concepteurs de programmes adaptatifs, les chercheurs en éducation et les sociologues soulignent l'importance des principes suivants : établir en classe un climat d'apprentissage qui incite les élèves à faire des efforts; organiser le programme; enseigner selon des objectifs clairement conçus et s'assurer que le temps en classe, du temps précieux, soit consacré à l'enseignement et aux efforts d'apprentissage des élèves.

Ce qu'il faut, dans la formation des enseignants avant l'emploi, c'est un examen du programme et des processus du point de vue de la préparation des enseignants à accomplir des tâches qui, réunies, définissent leurs responsabilités professionnelles et se traduisent par la création dans les classes d'un climat favorisant l'apprentissage des élèves. Les programmes habituels de formation des enseignants ne parviennent pas à répondre à ces buts. Dans un autre article (Ryan, 1984), nous avons procédé à une révision complète des programmes traditionnels de formation et de perfectionnement des enseignants. En résumé, nous avons d'abord découvert que les cours sur les théories et les méthodes sont compartimentés, ce qui fait que les candidats à l'enseignement arrivent mal à voir comment les cours théoriques peuvent les aider à prendre des décisions ou à faire leur travail. Si l'on concevait et dispensait ces cours en mettant l'accent sur les types de tâches des enseignants dont ils traitent effectivement, le lien serait évident. De plus, les connaissances issues de recherches seraient présentées en référence à leur utilité à aider les enseignants dans leurs responsabilités professionnelles de tous les jours.

Deuxièmement, nous avons décelé d'importantes omissions dans les programmes de formation des enseignants en cours d'emploi, en ce qui concerne plusieurs des tâches des enseignants. En particulier, on accorde peu d'attention aux tâches exigeant différents types de planification préactive par les enseignants. Les cours théoriques, par exemple, traitent rarement du caractère distinctif de la classe comme système écologique (Doyle, 1977) ou comme cadre de direction (Johnson et Brooks, 1979), de sorte que les futurs enseignants n'apprennent guère à comprendre les bases sur lesquelles ils devront prendre des décisions concer-

nant, d'une part, l'organisation d'un groupe particulier d'élèves en vue de leur instruction et, d'autre part, le choix d'un ensemble de règles et d'habitudes de comportements des élèves (Calfee et Brown, 1979). On enseigne la planification du programme en mettant l'accent de façon particulière sur les leçons individuelles en classe (Romberg et Price, 1983), probablement parce que l'enseignement lui-même est défini de façon restreinte comme étant des rencontres avec des élèves (Goodlad, 1983). On n'offre peu ou pas de formation spéciale sur le développement ou le choix de matériel de soutien pédagogique judicieux (Courter et Ward, 1983) ou sur le choix ou l'élaboration de devoirs et de tests appropriés pour les élèves ou même sur les décisions à prendre en ce qui concerne le niveau de rendement auquel les enseignants peuvent s'attendre de leurs élèves (Stones, 1984; Ryan et Schmidt, 1979). On enseigne rarement les stratégies de prise de décisions (Borko et al., 1979).

En somme, on suppose que les futurs enseignants pourront apprendre ce qu'ils doivent savoir en devenant des «apprentis» confiés à des enseignants expérimentés. L'accent est mis sur les comportements de ce «maître» enseignant, dans des situations interactives en classe. Cette approche «apprenti-maître» produit des tentatives de copier des «parcelles» de comportements d'enseignement, dont l'efficacité peut très bien dépendre du fait qu'ils font partie d'un ensemble global de comportements : «qui, s'il est morcelé et adopté par un autre, peut devenir inefficace et même nuisible» (Stones, 1984). Good, Grouws et Ebmeier (1983) ont renforcé notre point que les comportements d'enseignement individuels ne constituent pas une bonne façon de conceptualiser les comportements éducatifs. Les comportements en classe tirent leur signification des intentions des enseignants et de leurs plans préactifs. L'accomplissement par les enseignants de tâches préactives de planification importe plus que les comportements particuliers dont ils font montre en classe car ces comportements ne sont que des moyens de réaliser des plans.

Les genres d'implications qui accompagnent l'opinion que l'enseignement est un ensemble de tâches dans l'examen critique et la reformulation des programmes de formation des enseignants établissent un parallèle avec ses implications pour la supervision et l'évaluation des enseignants. Ailleurs (Ryan, 1985a), nous avons traité de ces implications de façon exhaustive. En somme, nous avons trouvé que les pratiques de supervision actuelles se caractérisent par une orientation «processus» de la gestion de l'enseignement, qu'elles emploient des jugements subjectifs sur l'efficacité des enseignants, jugements fondés en grande partie sur des observations des comportements en classe des enseignants et qu'elles

produisent des prescriptions ou des directives de comportements en classe («comment»).

En utilisant la conceptualisation des tâches des enseignants (illustration 1), nous avons soutenu que la supervision des enseignants devrait se réorienter vers une aide et une évaluation relatives aux tâches de planification préactives («quoi») auxquelles devraient se livrer les enseignants. La revue de la documentation précédant notre explication sur l'illustration 1 a révélé que les enseignants dérogeaient rarement à leurs plans, même lorsque la leçon se déroulait mal (Peterson et Clark, 1978). Autrement dit, les enseignants accomplissent ce qu'ils planifient. De plus, les très importantes tâches de planification des enseignants ne peuvent être observées, et partant, évaluées au cours de visites en classe. Des cinq unités structurales exigeant la réalisation de tâches par l'enseignant ou l'enseignante, une seule peut être directement observée en classe. Une des plus importantes «leçons» tirées de la recherche sur l'enseignement est que tout comportement en classe d'un enseignant particulier ou d'une enseignante particulière peut être ou ne pas être efficace, à n'importe quel moment, et dépend du pourquoi de son utilisation par l'enseignant ou l'enseignante et du caractère approprié de cette utilisation à ce moment dans cette classe. Si les enseignants accomplissent leurs tâches de planification et s'ils profitent des renseignements sur leurs élèves qu'ils obtiennent continuellement grâce à leurs tâches d'évaluation, ils sont alors en meilleure position que les superviseurs pour juger du bien fondé de leurs propres comportements. Essentiellement, les enseignants peuvent placer leurs comportements en contexte.

Nous avons également soutenu que les superviseurs des enseignants devraient non seulement cesser de concevoir l'enseignement strictement comme des rencontres avec les élèves, mais qu'ils devraient en outre cesser de concevoir la supervision des enseignants exclusivement comme l'observation et la rencontre d'enseignants individuels. Si on réoriente les efforts d'amélioration de l'enseignement dans les écoles de façon à aider les enseignants à accomplir leurs tâches préactives, les enseignants pourront travailler en équipe sur ces tâches, avec du soutien organisationnel (Howey et Vaughan, 1983). Nous avons noté que les enseignants individuels manquaient de temps et d'habiletés pour accomplir les tâches de planification nécessaires. Par conséquent, la supervision individuelle devrait compléter ce but majeur : offrir au niveau des écoles et des districts des initiatives et des programmes de perfectionnement du personnel reliés aux tâches de soutien et de développement éducatifs.

L'utilisation des tâches des enseignants pour faire le pont entre la recherche et la pratique

La conception (fonctionnelle) que l'enseignement est un ensemble de tâches donne naissance à un modèle d'organisation permettant d'examiner les connaissances issues de la recherche et de les utiliser pour améliorer la pratique. Comme nous l'avons déjà fait remarquer, en considérant les tâches discrètes (mais non liées entre elles) des enseignants, nous obtenons une façon de donner régulièrement aux enseignants une appréciation de ce qu'ils accomplissent. Il n'est pas essentiel que leur travail soit évalué seulement par les résultats des élèves aux tests. On a beaucoup à apprendre des recherches embrassant l'approche «comportements», mais à la condition que ces connaissances soient mises en perspective, c'est à dire que certains comportements sont probablement plus efficaces que d'autres pour aider les enseignants à accomplir une tâche particulière. De plus, certains «traits» des enseignants concourent plus que d'autres à la réalisation de certaines tâches (e.g. il y a des chances que les enseignants sympathiques et bienveillants établissent plus facilement de bons rapports avec les élèves). Ce qui importe, c'est que les «fins» d'un bon enseignement soient perçues comme la réalisation de tâches par les enseignants; les traits et les comportements sont susceptibles d'être les «moyens» pour parvenir à ces fins.

Dans cette perspective, nous voudrions illustrer comment les connaissances acquises par la recherche peuvent servir aux enseignants pour remplir d'importantes tâches ou fonctions. Dans son rôle de conseil auprès du département d'éducation de la Caroline du Sud, Anderson (1986a) a réduit à huit tâches primaires les tâches décrites dans notre illustration 1, puis il a décrit certains des éléments qu'il nous faut connaître sur la façon dont les enseignants les accomplissent et les «compensations» qu'ils en retirent. Puisque cet article n'est pas facilement disponible aux éducateurs canadiens, je le reproduis ici en grande partie, en espérant qu'il stimulera les discussions entre les enseignants, les administrateurs scolaires et les responsables de la formation des enseignants. Les quelques paragraphes suivants sont extraits de l'étude d'Anderson (1986a, pp.2-4).

Tâche 1. Planification de l'enseignement

La planification de l'enseignement aide les enseignants à atteindre plusieurs buts :

- Devenir conscients de la matière ou des objectifs importants, des méthodes indiquées pour évaluer à quel degré les élèves ont absorbé la matière ou maîtrisé les objectifs, et du matériel et des méthodes pédagogiques pertinents.
- Assurer la disponibilité du matériel de soutien pédagogique requis (e.g. lec-

tures supplémentaires, auxiliaires audio-visuels, logiciels), des devoirs et des tests.

- Évaluer le temps qu'il faudra prévoir pour les différents sujets, objectifs, chapitres des manuels ou activités (e.g. expériences en laboratoire).

- Faire correspondre les tests, évaluations, stratégies et méthodes d'enseignement avec le contenu et les objectifs du programme. Les enseignants sont capables d'orienter leur enseignement vers les objectifs d'apprentissage qu'ils ont fixés pour leurs élèves et évaluent l'apprentissage des élèves en conséquence.

- Faire correspondre le contenu et les objectifs des leçons et des unités plus grandes aux objectifs plus globaux du cours. Les élèves sont capables de voir le lien entre des faits isolés et le principaux concepts et généralisations qui «définissent» le sujet étudié.

- Concevoir ou établir les règles et les habitudes d'enseignement et de direction de classe.

La planification de l'enseignement s'effectue à divers niveaux. Tous les enseignants ne planifient pas aux mêmes niveaux; certains planifient au niveau du cours sur une base quotidienne, d'autres sur une base hebdomadaire, en reléguant la planification au niveau du cours au manuel et la planification quotidienne aux circonstances.

Les enseignants débutants et ceux qui ont de l'expérience planifient bien différemment. Les premiers tendent à préparer des plans de cours extrêmement précis et à les suivre fidèlement; ils sont un peu désemparés lorsque les circonstances les empêchent de suivre leurs plans. D'un autre côté, les enseignants d'expérience préparent des plans faisant davantage figure d'esquisses qui guident, mais ne déterminent pas ce qui se passe en classe. Si leurs plans se révèlent impraticables, qu'à cela ne tienne, ils ont plus d'un tour dans leur sac et improvisent. Enfin, les enseignants diffèrent dans leur approche globale de la planification (Clark et Yinger, 1979; 1980). Certains aiment planifier dans le détail et concevoir des plans qui guideront leur enseignement en fonction de leurs prévisions sur la façon dont les choses vont ou ne vont pas aller en classe. D'autres planifient par petites étapes, en analysant l'information au jour le jour, puis en décidant des actions à prendre subséquemment. Les bons enseignants peuvent aussi différer dans la façon dont ils planifient et dans la forme que prennent leurs plans, mais ils ont tous tendance à accomplir la tâche de la planification de l'enseignement.

Tâche 2. L'évaluation des élèves

Les enseignants évaluent leurs élèves afin de connaître leur niveau de connais-

sances, leurs aptitudes, attitudes et valeurs. Airasian (1984) a identifié trois types d'évaluations auxquelles ont fréquemment recours les enseignants : les évaluations «approximatives» au début de l'année ou du trimestre leur permettent de se faire une idée des connaissances au niveau du groupe et des individus. Les évaluations éducatives montrent comment les élèves ont profité (ou sont susceptibles de profiter) de l'enseignement tel que planifié et donné. Les évaluations éducatives peuvent comprendre une interrogation orale en classe, des devoirs ou des tests préparés par les enseignants. Les évaluations formelles sont celles qu'exigent la loi, l'état ou l'instance locale.

Les enseignants peuvent utiliser les évaluations pour prendre diverses décisions d'enseignement :

- Quand aborder une matière nouvelle ou de nouveaux objectifs.
- Accorder des notes aux élèves, déterminer les élèves capables de passer au niveau suivant et ceux devant être recalés.
- Choisir le contenu et les objectifs convenant bien aux élèves. Ces évaluations cherchent à établir le degré d'acquisition des prérequis cognitifs nécessaires qui permettront aux élèves d'apprendre de la matière ou des objectifs particuliers.
- Dépister quels élèves (ou sous-groupes d'élèves) auront besoin d'aide supplémentaire, ce qui exigera plus d'efforts de la part de l'enseignant ou enseignante et des arrangements dans le cas des élèves n'ayant pas besoin de cette aide.

De toutes façons, ce qui compte, ce n'est pas tant la forme de ces évaluations que l'information qu'elles procurent.

Tâche 3. Clarifier les règles et les habitudes de comportements

Clarifier les règles et les habitudes de comportements résulte en des comportements convenables chez les élèves de la classe. Quand il se produit des situations habituelles, comme passer des feuilles, demander la parole en classe et autres, les élèves n'adoptent pas un mauvais comportement et réagissent presque automatiquement.

Les règles et les habitudes sont apparentées de si près qu'il arrive souvent qu'on les confonde. Dit simplement, les règles sont des déclarations de ce qui n'est pas permis; ce sont des interdictions ou des contraintes sur les comportements des élèves en classe. D'un autre côté, les habitudes sont des modèles communs de comportements acceptables. Les habitudes donnent la façon d'accomplir efficacement les activités régulières de la classe. Les règles définissent

les comportements indésirables et les habitudes maintiennent le rythme des activités de la classe.

Plusieurs recherches ont révélé au moins trois caractéristiques des règles et des habitudes qui sont maintenant bien connues et acceptées.

- Les règles ne doivent pas être trop nombreuses et doivent être applicables de façon continue.
- Les habitudes importantes doivent être fixées tôt au cours de l'année scolaire ou au moment où une situation se présente.
- Il faut entretenir et renforcer les règles et les habitudes tout au long de l'année scolaire si on veut qu'elles demeurent efficaces.

Tâche 4. L'organisation de la classe

Une bonne organisation de la classe offre deux compensations reliées :

- La plus grande partie du temps de classe se passe en enseignement et en apprentissage.
- La forme d'enseignement (e.g. cours magistral, récitation, travail avec les médias, travail individuel) et les regroupements (e.g. classe entière, petits groupes, chacun pour soi) soutiennent les buts généraux du cours et le niveau d'avancement des élèves. On augmente ainsi les chances que les élèves apprennent ce qu'ils sont censés apprendre.

Les regroupements «en classe» sont un exemple particulier. Les différences individuelles dans le rendement des élèves de l'école secondaire sont beaucoup plus marquées que dans le cas des écoliers du palier élémentaire. On risque fort de retrouver de plus grosses différences dans les classes du secondaire dites homogènes que dans les classes hétérogènes (ou «non regroupées) du primaire. On peut résoudre ce problème en opérant des regroupements flexibles à l'intérieur des classes.

Comme on peut s'y attendre, les regroupements «en classe» causent bien des problèmes à l'enseignant ou l'enseignante et ces problèmes nécessitent une bonne direction de classe. Heureusement, la recherche récente propose des solutions. Les enseignants doivent :

- Donner des devoirs clairs aux groupes avec lesquels ils n'ont pas d'interaction directe.
- Créer dans ces groupes ce que Kounin et Sherman (1979) appellent un «pouvoir de cohésion» -- les élèves continuent à travailler, même en l'absence physique ou psychologique de l'enseignant ou enseignante.

- Prendre conscience de ce que Barr (1975) appelle la «taille du reste» -- le nombre d'élèves ne faisant pas partie du groupe en interaction avec l'enseignant ou l'enseignante. En général, plus la taille du reste est grande, plus on se retrouve avec des problèmes de direction de classe.

- Songer à la faisabilité et la désirabilité de ce qui a été appelé «l'apprentissage coopératif» (Johnson, Maruyama, et al., 1981), ou «l'apprentissage en équipe» (Slavin et Karweit, 1981). Dans l'apprentissage coopératif ou en équipe, un groupe d'élèves travaillent ensemble pour atteindre une tâche ou un but commun.

Tâche 5. La création d'un cadre d'apprentissage

Cette tâche est connue sous plusieurs appellations, dont la plus descriptive est celle-ci : «créer un cadre d'apprentissage» (Hunter, 1983). Lorsqu'on crée un cadre d'apprentissage, on se retrouve avec des élèves disposés à apprendre. Ici, cela veut dire que les élèves possèdent les prérequis cognitifs et affectifs pour apprendre.

- Les élèves sont au courant du contenu particulier ou des objectifs qu'ils ont à atteindre.

- Les liens entre la nouvelle matière ou les nouveaux objectifs et les apprentissages passés, actuels ou à venir sont clairs.

- Émotionnellement, les élèves sont prêts à l'apprentissage. Ils ont vraiment envie d'apprendre ou alors, on peut les obliger à fournir les efforts requis pour apprendre.

La tâche de créer un cadre d'apprentissage est particulièrement importante lorsque les enseignants font face à des apprenants peu préparés ou peu coopératifs. Les dernières recherches montrent que les enseignants peuvent créer un cadre (1) en procurant aux élèves les prérequis cognitifs nécessaires juste avant de leur enseigner la nouvelle matière ou de poursuivre de nouveaux objectifs, (2) en communiquant clairement leurs attentes sur ce que les élèves doivent apprendre et comment ils doivent manifester leur apprentissage, et (3) en utilisant des stimulants, le renforcement et les autres principes de la psychologie du comportement.

Tâche 6. L'enseignement par objectifs

Quand cette tâche est accomplie, les élèves apprennent précisément ce qu'ils sont censés apprendre. Les enseignants qui enseignent par objectifs ont en tête certains buts pédagogiques en disant aux élèves ce qu'ils leur disent, en leur montrant ce qu'ils leur montrent ou en leur donnant les devoirs qu'ils leur don-

nent. Ces enseignants ne vont pas faire lire Macbeth à leurs élèves simplement parce que tous devraient lire cette oeuvre classique, mais plutôt dans l'espoir que les élèves apprendront des choses sur le déroulement d'une intrigue ou les tourments des personnages. L'enseignement par objectifs exige une présentation initiale de la matière et une deuxième présentation si cela est nécessaire. En fait, l'importance de la deuxième présentation est cristallisée dans la tâche de l'enseignement par objectifs.

L'enseignement direct et la clarté d'expression sont les deux principales techniques de l'enseignement par objectifs. L'enseignement direct se produit quand l'enseignant ou l'enseignante prend activement part à la transmission aux élèves de connaissances ou d'habiletés clairement définies. La clarté de présentation signifie que cette transmission de connaissances et d'habiletés s'effectue de telle manière que la grande majorité des élèves saisissent et apprennent. Rosenshine (1983) offre plusieurs suggestions dérivées empiriquement pour aider les enseignants à enseigner par objectifs. Les enseignants devraient :

- Faire la démonstration de l'apprentissage visé à leurs élèves (lorsque cette fonction est indiquée).
- Organiser la matière et la présenter point par point.
- Donner des détails et des explications redondantes sur différents points.
- Utiliser plusieurs exemples variés et précis pour illustrer les points importants.
- Vérifier si les élèves ont compris avant de passer d'un point important au suivant, les interroger de façon serrée lorsque cela est indiqué et nécessaire.
- Insister sur les points importants jusqu'à ce que tous les élèves ou presque les aient compris.

Tâche 7. Prévoir de la pratique pour les élèves

La pratique offre deux compensations immédiates. Les élèves se souviennent plus rapidement des éléments appris et les utilisent plus rapidement, et ils retiennent ce qu'ils ont appris plus longtemps. Les exercices et les activités pratiques ont également une fonction d'évaluation. Les enseignants voient à quel point les élèves ont profité de la présentation initiale de la matière et identifient les élèves ayant besoin d'enseignement supplémentaire.

Les chercheurs ont identifié deux types de pratique. La pratique guidée s'effectue sous la supervision d'un enseignant ou d'une enseignante, d'un aide, d'un tuteur ou d'un camarade; elle dure habituellement peu longtemps et est destinée

à aider les élèves à réussir. La pratique autonome se fait sans la supervision directe de l'enseignant, d'un aide, d'un tuteur ou d'un camarade. Les devoirs à la maison sont un excellent exemple de pratique autonome. Le lien entre les deux types de pratique est assez clair : on devrait aider les élèves à obtenir un taux de succès raisonnable pendant des périodes de pratique guidée avant de leur donner l'occasion de s'exercer seuls. Rosenshine (1983) propose encore une fois plusieurs suggestions provenant de recherches, pour l'amélioration de la pratique guidée :

- Les élèves doivent connaître un taux élevé de succès.
- On devrait utiliser des indices aussi souvent que nécessaire pour guider les élèves vers les bonnes réponses.
- On doit exiger que les élèves expliquent leurs réponses (e.g. «Comment as-tu trouvé la réponse»?).
- Il faut corriger les erreurs au fur et à mesure qu'elles se produisent et ne pas les laisser s'accumuler.
- Si un nombre assez considérable d'élèves commet les mêmes erreurs, cela indique qu'il faudrait peut-être revenir sur cette matière.

Rosenshine émet également trois autres suggestions sur la pratique autonome : les directives pour la pratique autonome doivent être claires et concises; il faut vérifier si les élèves les ont bien comprises avant que l'exercice ne débute; les enseignants doivent surveiller cette pratique activement et offrir de l'aide si nécessaire.

Tâche 8. Maintenir la participation des élèves à l'apprentissage et punir les comportements indésirables

Même s'il semble que deux éléments réunis forment la huitième tâche des enseignants, ces éléments sont en réalité les deux côtés de la même médaille. La tâche première est de maintenir la participation des élèves à l'apprentissage, ou encore de tenir les élèves engagés dans l'apprentissage, à l'ouvrage. La deuxième tâche, punir les comportements indésirables, devient progressivement plus importante quand les élèves ne sont pas engagés dans l'apprentissage. Les élèves fournissent l'effort mental pour bien apprendre ce qu'ils sont censés apprendre.

Anderson (1981, 1984) a identifié une série de facteurs éducatifs associés à des niveaux supérieurs de temps de participation des élèves, ou de temps consacré à la tâche. Les élèves ont tendance à être plus engagés dans l'apprentissage et passer plus de temps à la tâche lorsque :

- Les tâches assignées présentent un niveau de difficulté approprié.
- Les tâches assignées sont intéressantes ou agréables à faire.
- Le rendement sur les tâches assignées est fréquemment surveillé et évalué.
- Les détails de «cuisine» du comportement en classe sont réduits au minimum.
- Les conditions physiques dans lesquelles l'apprentissage a lieu favorisent l'apprentissage.
- Les enseignants maintiennent le rythme des activités en classe.
- La curiosité des élèves est éveillée.
- Les enseignants communiquent clairement leurs attentes aux élèves.
- Les nouveaux apprentissages sont reliés aux précédents.
- L'accent est mis sur les aspects importants et pertinents du matériel et des activités pédagogiques.
- Les enseignants informent les élèves pour que ceux-ci sachent si leur rendement dans les tâches est simplement satisfaisant ou excellent.
- Le comportement orienté vers la tâche est renforcé. (Anderson, 1986a)

D'autres chercheurs ont également essayé d'organiser les connaissances issues de recherches pour qu'elles s'appliquent à la pratique de façon utile. Rosenshine (1983) a identifié les fonctions éducatives auxquelles ces connaissances s'appliquent. Stallings (1985) a choisi d'organiser ces applications de la recherche sur la base des caractéristiques des classes efficaces et de ce que les enseignants peuvent faire pour créer des conditions propices à l'apprentissage. Stallings a également identifié du matériel et des programmes disponibles pour aider les enseignants à développer un répertoire d'habiletés et de compétences appropriées. Dans l'espoir de stimuler l'intérêt pour l'article de Stallings, nous en donnons un exemple ici. Voici ce que Stallings (1985, p. 166) dit sur la direction de la classe :

> Il ne fait aucun doute que les classes bien dirigées s'en tirent mieux lors de tests de rendement scolaire (Brophy, 1981; Fisher et al., 1980; Good, 1979; Rosenshine et Berliner, 1978). «Comme les enseignants qui réussissent à bien diriger leurs classes maximisent le temps que les élèves consacrent à des activités scolaires, ceux-ci ont davantage de chances d'apprendre et cela se reflète par unrendement supérieur aux tests» (Brophy, 1982). Cependant, connaître ce fait n'aidera pas l'enseignant ou l'enseignante en formation avant l'emploi à apprendre comment bien diriger une classe. Des observations par les chercheurs Evertson, Anderson et Emmer, dans des écoles élémentaires et secondaires (1980), spécifiaient que la pratique peut être guidée dès les premiers jours de classe. Les chercheurs ont décrit en détail comment les enseignants efficaces établissaient et réalisaient leurs

> plans de direction, et par la suite, ils ont préparé à l'intention des enseignants une liste de points à vérifier. Les enseignants en formation avant l'emploi devraient tirer profit de cette documentation. (traduction libre)

En somme, le concept des tâches ou des fonctions des enseignants constitue un modèle d'organisation utile pour aborder l'efficacité des enseignants et les connaissances acquises grâce aux nombreuses recherches entreprises sur ce sujet. En réfléchissant sur sa tentative d'utiliser les fonctions comme heuristique d'organisation, Rosenshine (1983) a déclaré : «J'ai relu mes articles précédents pour voir si les sections correspondaient aux fonctions et j'ai été frappé de constater à quel point ces articles étaient fragmentés et de voir le nombre de considérations dont je n'avais pas tenu compte». Il a conclu qu'une liste de fonctions (ou tâches, dans notre terminologie) était un excellent modèle d'organisation préalable qui devrait être présenté aux enseignants avant qu'ils ne commencent des programmes précis de formation.

Résumé et conclusions

Dans cette monographie, nous avons brièvement passé en revue les différents types de recherches sur l'efficacité des enseignants auxquels se sont livrés les chercheurs sur une période de plus de quatre-vingts ans. À chaque génération de chercheurs, les études ont incorporé les connaissances accumulées précédemment et ont tenté de résoudre les problèmes de méthodologie et de faiblesses dans les fondements théoriques. Le plus sérieux problème des chercheurs a été et continue d'être l'établissement d'un point de départ, une façon de concevoir l'enseignement, qui rendrait vraiment compte de la complexité du métier d'enseignant. C'est seulement lorsque la recherche est dotée d'un point de départ valable que ses résultats donnent des directives utiles pour améliorer le rendement des enseignants.

Pendant la plus grande partie de ce siècle, la recherche sur l'enseignement s'est inspirée de la supposition générale que c'est ce que les enseignants sont qui les rend efficaces ou non. Les premières recherches en particulier se fondaient sur la supposition que les enseignants différaient selon qu'ils possédaient certains traits de personnalité typiques des enseignants efficaces. Par conséquent, la recherche s'attardait à définir ces traits «souhaitables», surtout en recueillant les jugements d'anciens élèves et d'éducateurs, puis en incorporant ces traits «souhaitables» dans des formules d'évaluation. Ces formules d'évaluation pouvaient être utilisées pour juger de l'efficacité que pourraient atteindre les candidats à l'enseignement et les enseignants déjà en

poste. Les administrateurs scolaires et les responsables de la formation des enseignants adoptèrent les suppositions sur les «bons» enseignants (i.e. échelles d'évaluation) et les «outils» d'observation qui étaient évidents dans la recherche à cette époque, même si rien ne prouvait que les enseignants qui possédaient ces traits étaient également ceux dont les élèves apprenaient le plus. Les tentatives pour démontrer ces liens (par les études présages-résultats) restaient pour la plupart sans succès et produisaient des résultats inconsistants.

À partir du milieu des années soixante, on vit apparaître une nouvelle approche à la recherche sur l'enseignement et on l'appela l'approche «processus-résultats». La supposition sous-jacente voulait que c'est ce que les enseignants font qui les rend efficaces ou pas. Plus précisément, cette recherche s'appuyait sur la supposition que les enseignants diffèrent dans leur utilisation des comportements éducatifs caractérisant les enseignants efficaces. Par conséquent, la recherche mettait l'accent sur l'identification des comportements «désirables», en grande partie par des systèmes d'observation à faible inférence et par des analyses des liens entre la fréquence moyenne d'utilisation des divers comportements des enseignants et de l'apprentissage moyen des élèves. Les tentatives visant à identifier des modèles de comportements «stables» des enseignants et à démontrer que ces modèles de comportements étaient reliés à l'apprentissage des élèves ont été de loin celles qui ont obtenu le plus de succès dans l'histoire de la recherche sur l'efficacité des enseignants. Les conclusions des études «processus-résultats» des années soixante-dix furent incorporées aux énoncés du genre «les enseignants devraient» et dans les programmes de formation des enseignants en fonction. Les listes de points à vérifier inspirées de l'approche sur les comportements et les observations en classe devinrent monnaie courante dans plusieurs districts scolaires.

C'est dans ce contexte que l'étude sur le contexte de la salle de classe de l'AIERS fut conçue comme une étude «processus-résultats» devant être répétée dans plusieurs pays autour du monde. Cependant, au moment même où l'on concevait l'étude, des critiques sur l'approche «processus-résultats» commençaient à apparaître dans la documentation. Ces critiques concernaient à la fois les suppositions que l'on faisait sur les processus d'enseignement et d'apprentissage et sur la méthodologie employée dans ces études. Il est juste de dire que les chercheurs qui ont contribué à la planification et la réalisation de l'étude de l'AIERS ont été davantage influencés par les publications proclamant que ce type de recherches était utile que par les critiques qui analysaient cette approche. Nos plans se déroulèrent comme prévu et nous avons découvert les faiblesses de l'approche «processus-résultats» par nos propres expériences.

À la suite de notre travail sur l'étude de l'AIERS et de notre examen attentif de la documentation, mes collègues et moi avons commencé à remettre en question la validité de notre conception que l'efficacité des enseignants repose sur leurs comportements et par conséquent, du paradigme de recherches «processus-résultats». Nous avons suggéré qu'un modèle d'efficacité des enseignants serait plus valide et utile s'il se fondait sur la conception que l'enseignement est un ensemble de tâches. Quand on conçoit l'enseignement comme un ensemble de tâches reliées entre elles, il y a une façon d'identifier aussi bien les omissions sérieuses dans la formation des enseignants que les erreurs importantes dans la façon de superviser les enseignants et de les aider à améliorer leurs pratiques d'enseignement. Les discussions dans la dernière partie de la monographie ont présenté une telle description des tâches des enseignants et ses implications pour les enseignants, les administrateurs scolaires et les responsables de la formation des enseignants. Le concept des tâches des enseignants fournit aussi un cadre organisationnel permettant de relier les conclusions des recherches à leurs applications pratiques. Nous avons en avons également traité dans la partie III.

Manifestement, la conception que l'enseignement est un ensemble de tâches rend obligatoires des changements à la conception et la réalisation des recherches futures sur l'efficacité des enseignants. Par exemple, l'attention sur les décisions préactives de planification des enseignants exige que l'on pense à de nouvelles méthodes pour recueillir l'information. On ne peut directement observer si et comment les enseignants accomplissent leurs tâches en visitant les salles de classe car ces tâches exigent de la part des enseignants du travail à l'extérieur de la salle de classe. L'accent mis sur la réalisation en classe de tâches d'enseignement et de direction exige aussi que l'on mette au point de nouvelles méthodes d'observation en classe. Ces nouvelles méthodes devraient graviter autour des événements au lieu d'être assujetties au temps, et les catégories de comportements qu'elles comprendraient devraient être organisées en fonction des tâches pertinentes (e.g. orienter, enseigner, évaluer, corriger, maintenir, discipliner). À cet égard, Anderson (1986b) a mis au point un système d'observation structurale nommé DePict (Describing Practices and Intentions of Classroom Teachers) qui est employé à la fois comme outil de recherche et comme partie du processus d'évaluation des enseignants en Caroline du Sud. On peut consulter les articles d'Anderson (1984; 1986a et b), de Ryan (1985b) et de Mandeville (1984) pour une discussion complète des implications pour les futures recherches.

En conclusion, lorsque nous considérons l'avenir, nous croyons que l'optimisme qui a débuté à la fin des années soixante-dix devrait se poursuivre. Grâce à la conception (fonctionnelle) que l'enseignement est un ensemble de tâches, nous disposons déjà de ce que Rosenshine (1983) appelle «d'abondantes richesses pour améliorer la pratique éducative». De plus, la tendance à la baisse de la clientèle scolaire semble être sur le point de se renverser, ce qui fait que les systèmes scolaires vont probablement se mettre à embaucher un grand nombre de nouveaux enseignants au cours des dix prochaines années (voir Bridges, 1986). Ces possibilités devraient donner lieu à des améliorations dans la qualité de l'enseignement dont profiteront les prochaines générations d'élèves, à la condition que nous commencions immédiatement à améliorer la formation des étudiants en pédagogie.

La formation des enseignants, aussi bien avant l'emploi qu'en cours d'emploi, devrait adopter les buts principaux du développement des compétences dont ont besoin les enseignants pour créer et entretenir un climat d'apprentissage axé sur les tâches à réaliser, faire participer les élèves à des activités reliées à l'apprentissage et avoir recours au genre d'enseignement utilisé, selon les recherches, par les enseignants efficaces. Nous sommes d'accord avec Medley (1979, p.26), quand il écrit :

> Personne ne trouve vraiment que nos enseignants efficaces sont... suffisamment efficaces; ce qu'il faut améliorer, c'est l'art d'enseigner lui-même. Même s'il est urgent d'améliorer les enseignants les moins efficaces, il importe aussi de perfectionner les plus efficaces. (traduction libre)

Mes collègues et moi croyons que la création d'un contexte de salle de classe orienté vers la réalisation de tâches et un usage productif du temps en classe sont des conditions favorisant un apprentissage efficace par les élèves. Nous avons tenté de suggérer quelques-unes des tâches importantes qui, si les enseignants les accomplissent, devraient les aider à établir ces conditions favorables et donc, à inciter les élèves à faire plus d'efforts et à mieux apprendre.

BIBLIOGRAPHIE

Airasian, P.W. Classroom Assessment and Educational Improvement. Portland, Oregon: Northwest Regional Educational Laboratory, 1984.

American Educational Research Association. Second Report of the Committee on Criteria of Teacher Effectiveness. Journal of Educational Research 46 (1953): 641-658.

Anderson, H. A Study of Certain Criteria of Teaching Effectiveness. Journal of Experimental Education 23 (1954): 41-71.

Anderson, L.M.; Evertson, C.M. & Brophy, J.E. An Experimental Study of Effective Teaching in First-Grade Reading Groups. Elementary School Journal 79 (1979): 193-223.

Anderson, L.W. The IEA Classroom Environment Study: Teaching for Learning. Comparative Education Review (sous presse).

Anderson, L.W. Research On Teaching and Educational Effectiveness. NASSP Curriculum Report 15 (1986a): 1-4.

Anderson, L.W. Understanding Teacher Behavior in the Classroom: A Must for Sound Evaluation. NASSP Bulletin (May 1986a): 42-50.

Anderson, L.W. Thoughts on Accomplished Teachers. The Palmetto Principal (Nov. 1985): 19-21.

Anderson, L.W. Attention, Tasks, and Time. In L.W. Anderson (Ed.), Time and School Learning: Theory, Research, and Practice. London: Croom Helm, 1984.

Anderson, L.W. Instruction and Time-On-Task: A Review. Journal of Curriculum Studies 13 (1981): 289-303.

Anderson, L.W. & Block, J.H. Mastery Learning. In D. Treffinger, J. Davis & R. Ripple (Eds.), Handbook on Teaching Educational Psychology. New York: Academic Press, 1976.

Anderson, L.W. & Jones, B. Designing Instructional Strategies Which Facilitate Learning for Mastery. Educational Psychologist 16 (1981): 121-138.

Averch, H.A.; Carroll, S.J.; Donaldson, T.S.; Kiesling, H.J. & Pincus, J. How Effective Is Schooling: Critical Review and Synthesis of Research Findings. Santa Monica, Calif.: Rand, 1972.

Barker, R. The Stream of Behavior As an Empirical Problem. In R. Barker (Ed.), The Stream of Behavior. New York: Appleton-Century Crofts, 1963.

Barr, A.S. The Validity of Certain Instruments Employed in the Measurement of Teaching Ability. In H.M. Walker (Ed.), The Measurement of Teaching Efficiency. New York: McMillan, 1935.

Barr, A.S. & Emans, L.M. What Qualities Are Prerequisite to Success in Teaching? Nation's Schools 6 (1930): 60-64.

Barr, R. How Children Are Taught to Read: Grouping and Pacing. School Review 83 (1975): 479-498.

Barr, R. & Dreeden, R. Instruction in Classrooms. In L.S. Shulman (Ed.), Review of Research in Education (Fifth Edition). Itasca, Ill.: F.E. Peacock, 1978.

Borko, H.; Cone, R.; Russo, N.A. & Shavelson, R.J. Teachers' Decision Making. In P. Peterson & H. Walberg (Eds.), Research on Teaching. Berkeley, Calif.: McCutchan, 1979.

Boyce, A.C. Methods for Measuring Teachers' Efficiency. In the Fourteenth Yearbook of the National Society for the Study of Education, Part II. Bloomington, Ill.: Public School Publishing Co., 1915.

Bridges, E.M. The Incompetent Teacher. Philadelphia: The Falmer Press, 1986.

Brophy, J. Classroom Organization and Management. Exposé présenté à la conférence sur la recherche en enseignement. Warrenton, Virg., 1982.

Brophy, J. Teacher Praise: A Functional Analysis. Review of Educational Research (1981): 5-32.

Brophy, J. & Evertson, C. Context Variables in Teaching. Educational Psychologist 12 (1978): 310-316.

Brophy, J.E. & Evertson, C.M. Process-Product Correlations in the Texas Teacher Effectiveness Study: Final Report. (Research Report 74-4). Austin: University of Texas, Research and Development Center for Teacher Education, 1974.

Burns, R.B. The Process and Content of Teaching: A Conceptual Framework. Evaluation in Education: An International Review Series 8 (1984): 95-112.

Calfee, R. & Brown, R. Grouping Students for Instruction. In D.L. Duke (Ed.), Classroom Management, Seventy-eighth Yearbook of the National Society for the Study of Education, Part II, Chicago: University of Chicago Press, 1979.

Charters, W.W. & Waples, D. The Commonwealth Teacher Training Study. Chicago: University of Chicago Press, 1929.

Clark, C. & Yinger, R. The Hidden World of Teaching: Implications of Research on Teacher Planning. Research Series No. 77, The Institute for Research on Teaching, Michigan State University, 1980.

Clark, C. & Yinger, R. Teachers' Thinking. In P. Peterson & H. Walberg (Eds.), Research on Teaching. Berkeley, Calif.: McCutchan, 1979.

Coleman, J.S.; Campbell, E.Q.; Hobson, C.J; McPartland, J.; Mood, A.M.; Weinfeld, F.D. & York, R.L. Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Comber, L.C. & Keeves, J.P. Science Education in Nineteen Countries: An Empirical Study, International Studies in Evaluation, Vol. 1. New York: John Wiley, 1973.

Courter, R.L. & Ward, B.A. Staff Development for School Improvement. In G.A. Griffin (Ed.), Staff Development, Eighty-second Yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1983.

Crawford, J. & Gage, N. L. Development of a Research-Based Teacher Education Program. California Journal of Teacher Education 4 (1977): 105-123.

Crawford, J.; Gage, N.L.; Corno, L.; Stayrook, N.; Mitman, A.; Schunk, D. & Stallings, J. An Experiment on Teacher Effectiveness and Parent-Assisted Instruction in the Third Grade. Stanford, Calif.: Center for Educational Research at Stanford, 1978.

Doyle, W. Paradigms for Research on Teacher Effectiveness. In L.S. Shulman (Ed.), Review of Research in Education (5th Edition). Itasca, Ill.: F.E. Peacock, 1978.

Doyle, W. Learning the Classroom Environment: An Ecological Analysis. Journal of Teacher Education 28 (1977): 51-55.

Elam, S.M. A Decade of Gallup Polls of Attitudes Toward Education 1969-1978. Bloomington, Indiana: Phi Delta Kappa, 1978.

Evertson, C.; Anderson, L.M. & Emmer, E. Effective Management at the Beginning of the School Year. Elementary School Journal 80 (1980): 219-231.

Evertson, C.; Anderson, L. M. & Emmer, E. Effective Management at the Beginning of the School Year in Junior High Classes. Austin: University of Texas, Research and Development Center for Teacher Education, 1980.

Fisher, C.W.; Berliner, D.C.; Filby, N.N.; Marliave, R.; Cohen, L.S. & Dishaw, M.M. Teaching Behaviors, Academic Learning Time and Student Achievement: An Overview. In C. Denham & A. Lieberman (Eds.), Time to Learn. Washington, D.C.: National Institute of Education, 1980.

Fordham, A. The Context of Teaching and Learning. Hawthorn, Victoria: Australian Council for Educational Research, 1983.

Gael, S. Job Analysis: A Guide to Assessing Work Activities. San Francisco: Jossey-Bass, 1983.

Gage, N.L. The Scientific Basis of the Art of Teaching. New York: Teachers College Press, 1978.

Gage, N.L. Research on Cognitive Aspects of Teaching. In The Way Teaching Is. Washington, D.C.: Association for Supervision and Curriculum Development and NEA Center for the Study of Instruction, 1966.

Gage, N.L. (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Gage, N.L. Discours publié dans "Proceedings," Research Resume, Vol. 16. Burlingame, Calif.: California Teachers Association, 1960.

Gage, N.L. & Fordham, A. A Proposal for an IEA Classroom Environment Study: Teaching for Learning. IEA Secretariat (University of Stockholm), 1980.

Good, T. Teacher Effectiveness in the Elementary School: What We Know About It Now. Journal of Teacher Education 6 (1979): 105-113.

Good, T.; Biddle, B. & Brophy, J. Teachers Make a Difference. New York: Holt, Rinehart & Winston, 1975.

Good, T.L.; Grouws, D.A. & Ebmeier, H. Active Mathematics Teaching. New York: Longman, 1983.

Good, T.L & Grouws, D. Missouri Mathematics Effectiveness Project: Experimental Study in Fourth Grade Classrooms. Journal of Educational Psychology 71 (1979): 355-362.

Goodlad, J.I. A Place Called School. New York: McGraw-Hill, 1983.

Gump, P. Environmental Guidance of the Classroom Behavioral System. In B. Biddle & W. Ellena (Eds.), Contemporary Research on Teacher Effectiveness. New-York: Holt, Rinehart & Winston, 1964.

Hanna, G. & Ryan, D.W. Profiles of Effective Teachers of Grade 8 Mathematics. Exposé présenté à la réunion annuelle de l'American Educational Research Association, Chicago, 1985.

Hart, J.W. Teachers and Teaching. New York: Macmillan, 1936.

Hathaway, J. & Lockheed M.E. Introduction: The Social Sciences and Education. In J. Hathaway & M.C. Lockheed (Eds.), The Contributions of the Social Sciences to Educational Policy and Practice: 1965-1985. Berkeley, Calif.: McCutchan, 1986.

Howey, K.R. & Vaughan, J.C. Current Patterns of Staff Development. In G.A. Griffin (Ed.), Staff Development, Eighty-second Yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1983.

Hunter, M. Mastery Teaching. El Segundo, Calif.: TIP Publications, 1983.

Johnson, D.; Maruyama, G.; Johnson, R.; Nelson, D. & Skon, L. Effects of Cooperative, Competitive and Individualistic Goal Structures on Achievement: A Meta-Analysis. Psychological Bulletin 89 (1981): 47-62.

Johnson, M. & Brooks, H. Conceptualizing Classroom Management. In D.L. Duke (Ed.), Classroom Management, Seventy-eighth Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1979.

Jones, R. The Prediction of Teaching Efficiency from Objective Measures. Journal of Experimental Education 15 (1946): 85-99.

Joyce, B. & Weil, M. Models of Teaching. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Kounin, J.S. & Sherman, L.W. School Environments as Behavior Settings. Theory into Practice 18 (1979): 145-149.

Kozol, J. Death at an Early Age. Boston: Houghton Mifflin, 1967.

Kratz, H.E. Characteristics of the Best Teachers as Recognized by Children. Pedagogical Summary 3 (1896): 413-418.

Levin, T. & Lewy, A. Classroom Environment Study Research Proposal. IEA Secretariat (University of Stockholm), 1978.

Livingstone, D.W.; Hart, D.J. & Davie, L.E. Public Attitudes Toward Education in Ontario. Toronto: OISE Press, 1985.

Mandeville, G.K. Reanalyzing Teaching Research Data: Promises and Pitfalls. Evaluation in Education: An International Review Series 8 (1984): 153-166.

McCall, W. & Krause, G. Measurement of Teacher Merit. Journal of Educational Research 53 (1959): 73-5.

McDonald, F.J. & Elias, P. The Effects of Teaching Performance on Pupil Learning, Beginning Teacher Evaluation Study: Phase II, Final Report, Vol. 1. Princeton, N.J.: Educational Testing Service, 1976.

Medley, D.M. Teacher Effectiveness. In H.E. Mitzel (Ed.), Encyclopedia of Educational Research (5th Ed.). New York: The Free Press, 1982.

Medley, D.M. The Effectiveness of Teachers. In P. Peterson & H. Walbert (Eds.), Research On Teaching. Berkeley, Calif.: McCutchan, 1979.

Medley, D.M. Teacher Competence and Teacher Effectiveness. Washington, D.C.: American Association of Colleges for Teacher Education, 1977.

Mitzel, H.E. Teacher Effectiveness. In C.W. Harris (Ed.), Encyclopedia of Educational Research (3rd Ed.) New York: Macmillan, 1960.

Peterson, P. & Clark, C. Teachers' Reports of Their Cognitive Processes During Teaching. American Educational Research Journal 15 (1978): 555-565.

Purves, A.C. The IEA Empire: A Memoire. Comparative Education Review (sous presse).

Purves, A.C. Literature Education in Ten Countries: An Empirical Study, International Studies in Evaluation, Vol. 2. New York: John Wiley, 1973.

Romberg, T.A. & Price, G.G. Curriculum Implementation and Staff Development as Cultural Change. In G.A. Griffin (Ed.), Staff Development, Eighty-second Yearbook of the National Society for the Study of Education, Part II. Chicago: University of Chicago Press, 1983.

Rosenshine, B. Teaching Functions in Instructional Programs. The Elementary School Journal 83 (1983): 335-351.

Rosenshine, B. Content, Time, and Direct Instruction. In P. Peterson & H. Walberg (Eds.), Research on Teaching. Berkeley, Calif.: McCutchan, 1979.

Rosenshine, B. Classroom Instruction. In N.L. Gage (Ed.) The Psychology of Teaching Methods, Seventy-fifth Yearbook of the National Society for the Study of Education, Part I. Chicago: University of Chicago Press, 1976.

Rosenshine, B. Teaching Behaviors and Student Achievement. London: National Foundation for Educational Research, 1971.

Rosenshine, B. The Stability of Teacher Effects upon Student Achievement. Review of Educational Research 40 (1970): 647-662.

Rosenshine, B. & Berliner, D.C. Academic Engaged Time. British Journal of Teacher Education 4 (1978) 3-16.

Ryan, D.W. Preactive and Proactive Supervision of Mastery Learning Programs. In D.U. Levine (Ed.), Improving Student Achievement through Mastery Learning Programs. San Francisco: Jossey-Bass, 1985a.

Ryan, D.W. Research on Teaching and Teacher Education: Changing Views on Desirable Research Design and Method. In I.E. Housego & P.P. Grimmett (Eds.), Teaching and Teacher Education: Generating and Utilizing Valid Knowledge for Professional Socialization. Vancouver: Université de Colombie-Britannique, Faculté d'Éducation, 1985b.

Ryan. D. W. Improving the Content and Process of Pre- and In-Service Teacher Training. Evaluation in Education: An International Review Series 8 (1984): 113-131.

Ryan, D.W. Classroom Observation Strategies. Exposé présenté à la conférence sur l'amélioration de l'enseignement, Université Laurentienne, 1981.

Ryan, D.W. & Anderson, L.W. (Eds.), Rethinking Research on Teaching: Lessons Learned from an International Study, numéro spécial de Evaluation in Education: An International Review Series 8 (1984).

Ryan, D.W. & Anderson, L.W. The IEA Classroom Environment Study: Promises and Pitfalls. Evaluation in Education: An International Review Series 8 (1984): 87-93.

Ryan, D.W.; Hildyard, A.; Jantzi, D.; Pike, R.; Hanna, G. & Quinlan, F. The Ontario IEA Classroom Environment Study. Rapport final au ministère de l'Éducation de l'Ontario, 1984.

Ryan, D.W. & Schmidt, M. Mastery Learning: Theory, Research and Implementation. Toronto: Ministère de l'Éducation de l'Ontario, 1979.

Silberman, C. Crisis in the Classroom: The Remaking of American Education. New York: Random House, 1970.

Slavin, R.E. & Karweit, N. Cognitive and Affective Outcomes of an Intensive Study Team Learning Experience. Journal of Experimental Education 50 (1981): 29-35

Smith, L. & Geoffrey, W. The Complexities of an Urban Classroom. New York: Holt, Rinehart & Winston, 1968.

Soar, R.S. Final Report, Follow Through Classroom Process Measurement and Pupil Growth (1970-1971). Gainesville: Institute for the Development of Human Resources, University of Florida, 1973.

Soar, R.S. Teacher-Pupil Interaction. In J.R. Squire (Ed.), A New Look at Progressive Education. Washington, D.C.: Association for Supervision and Curriculum Development, 1972.

Stallings, J. Implications from the Research on Teaching for Teacher Preparation. In I.E. Housego & P.P. Grimmett (Eds.), Teaching and Teacher Education: Generating and Utilizing Valid Knowledge for Professional Socialization. Vancouver: Université de Colombie-Britannique, Faculté d'éducation, 1985.

Stallings, J.; Needels, M. & Stayrook, N. How to Change the Process of Teaching Basic Reading Skills in Secondary Schools: Phase II and Phase III; Final Report. Menlo Park, Calif.: SRI International, 1979.

Stallings, J. & Kaskowitz, D. Follow Through Classroom Observation Evaluation, 1972-1973 (SRI Project URU-7370). Menlo Park, Calif.: SRI International, 1974.

Stephens, J. The Process of Schooling. New York: Holt, Rinehart & Winston, 1967.

Stones, E. Practice Teaching. In T. Husen & T.N. Postlethwaite (Eds.), International Encyclopedia of Education. Oxford: Pergamon Press, 1984.

Thorndike, R.L. Reading Comprehension Education in Fifteen Countries, International Studies in Evaluation, Vol. 3. New York: John Wiley, 1973.

Veldman, D. & Brophy, J. Measuring Teacher Effects on Pupil Achievement. Journal of Educational Psychology 66 (1974): 319-324.

Walberg, H.J. A Nation at Risk: Retrospect and Prospect. In T.M. Tomlinson & H. Walberg (Eds.), Academic Work and Educational Excellence. Berkeley, Calif.: McCutchan, 1986.

ANNEXE A

CATÉGORIES, VARIABLES ET INSTRUMENTS

Catégories/Variables	Instruments
CARACTÉRISTIQUES DE LA COLLECTIVITÉ	
Emplacement de l'école	QDE
CARACTÉRISTIQUES DE L'ÉCOLE	
Type d'école (TYPEÉ)	QDE
Organisation des niveaux scolaires	QDE
Taille de l'école	QDE
Temps alloué au sujet (par année)	QDE
CARACTÉRISTIQUES DES ENSEIGNANTS	
Sexe (SEXEN)	QDE
Âge (ÂGEEN)	QDE
Années d'expérience (ANS)	QDE
Brevet	QDE
Matière de spécialisation	QDE
Nombre d'heures de contact avec la matière (par semaine)	QDE
Nombre d'autres matières enseignées	QDE
Total des heures de contact (par semaine)	QDE
CARACTÉRISTIQUES FAMILIALES	
Scolarisation du père	QDÉ
Scolarisation de la mère	QDÉ
Profession du père	QDÉ
Langue parlée à la maison	QDÉ
CARACTÉRISTIQUES PERSONNELLES DE L'ÉLÈVE	
Sexe (SEXÉL)	QDÉ
Âge (ÂGÉL)	QDÉ
Prétest (PRÉTEST)	PRÉTEST
Aspirations (ASPIRATION)	QDÉ

Attitude envers l'école (ATTIÉC)	QDÉ
Attitude envers la matière (PRÉATTIMAT)	QDÉ
CONTEXTE DE LA SALLE DE CLASSE	
Utilisation typique du regroupement	QDE
Utilisation observée du regroupement	ICM
Utilisation des types de matériels	QDE
Accès des élèves à des livres en dehors de l'école	QDE
Accent mis sur les leçons	EPO
Objectifs des leçons	EPO
Taille de la classe	FRI
Nombre d'adultes dans la classe	FRI
Assiduité des élèves	FRI
PERCEPTIONS DES ENSEIGNANTS	
Autonomie des enseignants	QDE
Niveau relatif d'habiletés de la classe	QDE
Pourcentage d'élèves ayant besoin d'enseignement correctif	QDE
QUANTITÉ D'APPRENTISSAGE	
Quantité de devoirs	QDE
Temps pour les devoirs	QDE
Chances d'apprentissage	CA
Temps alloué	FRI
Temps d'enseignement	ICM
ORIENTATION DES ÉLÈVES	
Précision des objectifs	QDE
Révision de la matière	INST
L'ENSEIGNEMENT AUX ÉLÈVES	
Interactions d'enseignement	INST
Surveillance par l'enseignant ou l'enseignante	INST
Surveillance du travail individuel	QDE
Surveillance du travail en groupes	QDE
Enseignement magistral	INST

Discours et discussions	INST
Exercices oraux et répétitions	INST
Lecture individuelle	INST
Travail écrit individuel	INST
Travail en laboratoire	INST
Sens des interactions verbales	ICM
Explications	ICM
Explications à l'aide des matériels	ICM
Démonstration	ICM
Utilisation d'exemples	ICM
Indices de structure	ICM
Directives	ICM
Questions des élèves	ICM
Interventions des élèves	ICM
Interactions sociales	ICM
Silence (absence d'interactions verbales)	ICM
Fréquence de l'aide supplémentaire	QFE
Nombre d'élèves recevant de l'aide supplémentaire	QFE
Caractère opportun de l'aide supplémentaire	QFE
Forme de l'aide supplémentaire	QFE
Type de l'aide supplémentaire	QFE

L'ÉVALUATION DES ÉLÈVES

Questions à des élèves	ICM
Questions à l'ensemble du groupe	ICM
Questions faisant appel à la mémoire des élèves	ICM
Questions d'ordre supérieur	ICM
Questions d'opinion	ICM
Questions d'efficacité	ICM
Quantité de questions	QFE
Attend les réactions de l'élève ou des élèves	ICM
Brève réaction d'un ou des élèves	ICM
Récitation par les élèves	ICM
Réaction prolongée d'un ou des élèves	ICM
L'élève ne répond pas	ICM
Acquiescement positif	ICM

Dit qu'une réponse n'est pas la bonne	ICM
Répète la question	ICM
Redirige la question	ICM
Donne la réponse	ICM
Interroge de façon soutenue	ICM
Interroge un élève de façon soutenue	ICM
Interroge le groupe de façon soutenue	ICM
Tests aux fins de notation	QFE
Tests de diagnostic	QFE
DIRECTION DES ÉLÈVES	
Gestion des règles et des habitudes	QFE, INST
Interactions sur les règles	ICM
Interactions transitionnelles	ICM
Actions disciplinaires	QFE, ICM
Non intervention de la part de l'enseignant	INST
PERCEPTIONS DES ÉLÈVES	
Orientation vers les tâches (TÂCHES)	QFÉ
Structure des leçons (STRUC)	QFÉ
Réactions de la part de l'enseignant (RÉACT)	QFÉ
PARTICIPATION DES ÉLÈVES	
Pourcentage des élèves participant à l'apprentissage	RAI, INST
RÉSULTATS DES ÉLÈVES	
Rendement des élèves (POSTTEST)	POSTTEST
Attitude envers la matière (POSTATTIMAT)	QFÉ